ROBUST PORTFOLIO OPTIMIZATION WITH RISK MEASURES UNDER DISTRIBUTIONAL UNCERTAINTY

A DISSERTATION SUBMITTED TO

THE GRADUATE SCHOOL OF ENGINEERING AND SCIENCE

OF BILKENT UNIVERSITY

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR

THE DEGREE OF

DOCTOR OF PHILOSOPHY

IN

INDUSTRIAL ENGINEERING

By

A. Burak Paç

July 2016

ROBUST PORTFOLIO OPTIMIZATION WITH RISK MEASURES UNDER DISTRIBUTIONAL UNCERTAINTY

By A. Burak Paç

July 2016

We certify that we have read this dissertation and that in our opinion it is fully adequate, in scope and in quality, as a dissertation for the degree of Doctor of Philosophy.

Mustafa Ç. Pınar(Advisor)

Savaş Dayanık

Banu Yüksel Özkaya

Kemal Yıldız

Gerhard-Wilhelm Weber

Approved for the Graduate School of Engineering and Science:

Levent Onural
Director of the Graduate School

ABSTRACT

ROBUST PORTFOLIO OPTIMIZATION WITH RISK MEASURES UNDER DISTRIBUTIONAL UNCERTAINTY

A. Burak Paç
Ph.D. in Industrial Engineering
Advisor: Mustafa Ç. Pınar
July 2016

In this study, we consider the portfolio selection problem with different risk measures and different perspectives regarding distributional uncertainty. First, we consider the problem of optimal portfolio choice using the first and second lower partial moment risk measures, for a market consisting of n risky assets and a riskless asset, with short positions allowed. We derive closed-form robust portfolio rules minimizing the worst case risk measure under uncertainty of the return distribution given the mean/covariance information. A criticism levelled against distributionally robust portfolios is sensitivity to uncertainties or estimation errors in the mean return data, i.e., Mean Return Ambiguity. Modeling ambiguity in mean return via an ellipsoidal set, we derive results for a setting with mean return and distributional uncertainty combined. Using the adjustable robustness paradigm we extend the single period results to multiple periods in discrete time, and derive closed-form dynamic portfolio policies.

Next, we consider the problem of optimal portfolio choice minimizing the Conditional Value-at-Risk (CVaR) and Value-at-Risk (VaR) measures under the minimum expected return constraint. We derive the optimal portfolio rules for the ellipsoidal mean return vector and distributional ambiguity setting. In the presence of a riskless asset, the robust CVaR and VaR measures, coupled with a minimum mean return constraint, yield simple, mean-variance efficient optimal portfolio rules. In a market without the riskless asset, we obtain a closed-form portfolio rule that generalizes earlier results, without a minimum mean return restriction.

In the final problem, we have a change of perspective regarding uncertainty. Rather than the information on first and second moments, knowledge of a nominal distribution of asset returns is assumed, and the actual distribution is considered to be within a ball around this nominal distribution. The metric choice on the probability space is the Kantorovich distance. We investigate convergence of the risky investment to uniform portfolio when a riskless asset is available. While uniform investment to risky assets becomes optimal, it is shown that as the uncertainty radius increases, the total allocation to risky assets diminishes. Hence, as uncertainty increases, the risk averse investor is driven out of the risky market.

Keywords: Robust Portfolio Choice, Distributional Robustness, Adjustable Robustness, Lower Partial Moments, Dynamic Portfolio Rules, Ellipsoidal Uncertainty, Conditional Value-at-Risk.

ÖZET

DAĞILIM BELİRSİZLİĞİ ALTINDA PORTFOLYO RİSK ÖLÇÜLERİNİN GÜRBÜZ OPTİMİZASYONU

A. Burak Paç
Endüstri Mühendisliği, Doktora
Tez Danışmanı: Prof. Mustafa Ç. Pınar
Temmuz 2016

Bu çalışmada, portfolyo seçimi problemi farklı risk ölçüleri ve dağılım belirsizliği açısından farklı bakış açıları altında ele alınmıştır. Öncelikle, n adet risk içeren ve bir tane risk içermeyen yatırım aracı bulunan bir pazar için, birinci ve ikinci dereceden alt kısmi moment risk ölçüleri kullanılarak en iyi portfolyo seçimi problemi düşünülmüştür. Risk içeren yatırım araçlarının getiri dağılımlarının belirsiz olduğu ve ortalama/eşdeğişirlik (kovaryans) bilgilerinin bilindiği varsayılarak, en kötü durum için risk ölçülerini eniyileyen, kapalı formül şeklinde gürbüz portfolyo seçimi kuralları elde edilmiştir. Dağılımsal olarak gürbüz portfolyo kuralları, belirsizliğe veya ortalama getiri verisindeki tahmin hatalarına (ortalama getiri bulanıklığı) karşı hassasiyetinden dolayı eleştirilir. Ortalama getiri verisindeki bulanıklık elipsoid şeklindeki bir küme ile modellenerek, ortalama getiri ve dağılımsal belirsizliğin bir arada olduğu bir durum için sonuçlar elde edilmiştir. Ayarlanabilir gürbüzlük paradigması kullanılarak, tek dönem için elde edilen sonuçlardan faydalanılarak, çok dönemli yatırımlar için kapalı formül şeklinde dinamik portfolyo kuralları elde edilmiştir.

Daha sonra, minimum beklenen getiri kısıtı altında Koşullu Risk Altındaki Değer (CVaR) ve Risk Altındaki Değer (VaR) ölçülerini enküçükleyen amaç fonksiyonu ile en iyi portfolyo seçimi problemi üzerine çalışılmıştır. Elipsoid şeklindeki ortalama getiri yöneyi ve dağılımsal bulanıklık için eniyi portfolyo kuralı elde edilmiştir. Pazarda risk içermeyen bir yatırım aracı bulunduğunda, minimum ortalama getiri kısıtı ile birlikte gürbüz CVaR ve VaR ölçüleri, basit, ortalama ve değişirlik bakımından etkin optimal portfolyo kurallarını doğurmuştur. Risk içermeyen bir yatırım aracı bulunan bir pazar için, minimum ortalama getiri kısıtı olmadan, daha önceki sonuçları genelleştiren kapalı formül şeklinde bir portfolyo kuralı elde edilmiştir.

Son problemimizde belirsizliğe bakış açısında bir değişiklik yapılmıştır. Birinci ve ikinci moment bilgileri yerine, yatırım araçlarının nominal dağılımı bilgisine sahip olunduğu ve gerçek dağılımın bu nominal dağılımın etrafındaki bir kürede olduğu varsayılmıştır. Ayrıca olasılık uzayının Kantorovich uzaklık metriğine sahip olduğu varsayılmıştır. Risk içermeyen bir yatırım aracı bulunduğunda, riskli yatırımın eşit dağılımlı portfolyoya yakınsadığı gözlemlenmiştir. Risk içeren yatırım araçları için eşit dağılımlı yatırım optimal iken, belirsizlik yarıçapı arttıkça riskli yatırım araçlarına toplam yatırımın azaldığı gösterilmiştir. Diğer bir deyişle, belirsizlik arttıkça, riskten kaçınan yatırımcının riskli pazardan çıktığı gŏsterilmiştir.

Anahtar sözcükler: Gürbüz Portfolyo Seçimi, Dağılımsal Gürbüzlük, Ayarlanabilir Gürbüzlük, Alt Kısmi Momentler, Dinamik Portfolyo Kuralları, Elipsoid Biçimli Belirsizlik, Koşullu Risk Altındaki Değer.

Acknowledgement

I would like to express my gratitude to my advisor Prof. Mustafa Pınar for his invaluable guidance, support, patience and encouragement during my graduate study. I am extremely lucky to have such a great advisor. It was a great experience of collaboration in research, as Prof. Pınar guided me through phases where he actively leaded research, beyond providing research directions, later taking me to a point where I was supposed to seek research directions on my own.

I am grateful to Asst. Prof. Banu Yüksel Özkaya and Assoc. Prof. Savaş Dayanık for devoting their valuable time to read each part of my work and providing precious suggestions. I am also grateful for their valuable contributions, their supportive and positive attitude in committee meetings. I also want to express my gratitude to Asst. Prof. Kemal Yıldız and Prof. Gerhard-Wilhelm Weber for accepting to be a member of my examination committee and for their valuable suggestions.

I would like to take this opportunity to express my gratitude to Prof. Ihsan Sabuncuoğlu. When I look back, I see his encouraging, supportive attitude with me at every difficult stage of graduate studies.

It was an honor to be a member of Bilkent University Department of Industrial Engineering, and I would like to thank each member of the department.

My friends and office mates Ece Demirci, Esra Koca, Gizem Özbaygın, Hatice Çalık, Merve Meraklı, Nihal Berktaş, Ramez Kian and Kamyar Kargar deserve special thanks for their friendship and support. Throughout the many years in Bilkent University, I enjoyed friendship with many people who have become dear to me. I want to thank each of them for making my life more beautiful and the joy of living they bring. I feel very lucky to have so many great people in my life.

Finally, I would like to express my deepest gratitude to my family for their endless love and trust.

Contents

List of Figures

Chapter 1

Introduction

Deterministic mathematical programs have fixed parameter values, and optimal solutions are obtained based on these fixed values of parameters. In reality, these parameters that are considered to have fixed values involve uncertainty due to measurement errors, noise, fluctuations caused by change of conditions. In many applications, these parameters are estimated based on historical data to forecast future realizations, and therefore might involve large drifts from their actual values at the time estimation data was collected. Often, the optimal solutions of these deterministic problems turn infeasible or suboptimal due to very small perturbations in the parameters. The effect of such perturbations is observed using sensitivity analysis techniques, which enters the scene after an optimal solution is found for the deterministic problem, and is a reactive approach -as named in Mulvey et al.[1]-, in that it assesses the reaction of the optimal solution to drifts in parameter values. This post-optimality technique does shed light on how the feasibility and optimality properties of the deterministic solution change with parameter drifts, and whether the optimal solution is stable in terms of feasibility and optimality. Indeed, even small drifts in parameters might have devastating consequences due to rendering the existing solution infeasible, or highly suboptimal after the drift. But only knowing the reaction of the optimal solution to the parameter drift is not of practical importance by itself. What is desirable is to find a solution of high performance while being immunized against the uncertainty in parameter values. Two methodologies, Robust Optimization and Stochastic Optimization, have different modeling perspectives for seeking high performing solutions that are immune to parameter uncertainties.

In Robust Optimization (RO) methodology, data uncertainty is modelled based on sets, and a selection of variables is required to satisfy model constraints for each parameter setting inside the uncertainty set, if it is to be called a (robust) solution of the RO problem. A robust optimal solution is a robust solution with the best worst-case performance among the robust solutions, i.e., the worst objective value it attains over entire parameter uncertainty set is better compared to those for other robust solutions. With the set based uncertainty introduced, the problem arising is called the robust counterpart (RC) of the deterministic mathematical program with fixed parameters.

Modeling uncertainty with sets provides great flexibility in representation of the region of parameter uncertainty, which increases applicability of RO on a broad variety of real world problems. While the structure of the uncertainty set can be tailored in conformance to many forms of parameter uncertainty, the size of the uncertainty set can be adjusted to achieve a balance between securing feasibility and obtaining a solution of high performance. The robust counterpart is in general computationally more complex than the deterministic problem, but with a careful selection of uncertainty set structure, in many cases it is possible to reformulate robust optimization models as mathematical program formulations known to be computationally tractable. Specifically, with ellipsoidal uncertainty sets it is possible to represent or approximate many forms of uncertainty. Moreover, mathematically desirable properties of ellipsoidal uncertainty sets, such as convexity, give rise to tractable formulations in many classes of mathematical programs/optimization problems.

In the Stochastic Optimization (SO) methodology, data uncertainty is modeled by probability distributions. For this to be possible, parameter uncertainty should be of stochastic nature, and probability distribution types for parameters should be known. Securing feasibility of solution for critical constraints can be achieved by probability chance constraints. In this case, the event that these critical constraints are not satisfied -even if it has a very small probability- should not have devastating consequences. Objective functions in SO models are generally based on moments. Typical objectives in stochastic portfolio selection are expected utility/expected return maximization, variance minimization or a maximization of a combination of expected return and variance penalized by a coefficient.

RO can be considered as the only strategy immunizing solutions against parameter uncertainty when uncertainty does not have probabilistic nature. Although SO might seem to be the methodology to implement otherwise, in many cases of stochastic uncertainty, it is more plausible to use RO models. Except for cases where distribution types can be inferred theoretically, as when Central Limit Theorem is applicable, it can be very difficult to determine probability distributions beyond simplifying by assumptions and using partial information such as moments of low degree. In most of the cases, the number of observations necessary to determine joint distribution of variables in multiple-dimensions renders stochastic modeling implausible. Even if the probability distributions of the parameters are known, RO can be preferable due to tractability reasons. While introduction of set based uncertainty might increase computational complexity for many deterministic mathematical programs, stochastic modeling of uncertainty in similar cases generally results in NP-hard problems. The mean-variance model in Markowitz [2], and mean-value-at-risk model in Benati and Rizzi [3] are examples of this. Computational complexities of scenario based multi-period SO models grow rapidly with number of scenarios and periods. Except for special cases, such as when chance constraints define convex sets, models involving probability chance constraints are generally intractable. In such cases where probability distribution information is available, but SO models are intractable, RO models can be preferred due to being practically solvable. With RO models, probability distribution information -when available- can be used to adjust uncertainty set sizes to provide desired probabilistic guarantees for the robust solution.

When the parameter data is known to be of probabilistic nature, but it is not plausible to find out the exact distribution type, or again, when robustness is preferable for tractability reasons, distributional robustness bridges the gap between RO and SO. In this case, parameters are considered to assume probability distributions from diverse families of distribution types. In this case, solution is immunized against the uncertainty in distribution, comparing solutions with respect to the worst cases arising over all distributions in the families describing the uncertainty. For many popular utility/disutility functions and risk measures, tight bounds have been derived for worst-case behavior over various distribution families. Plugging in these tight-binding functions instead of the supremums of functions or measures over families defining distribution uncertainty, elegant mathematical program formulations are obtained, often with tractability properties allowing for practical solvability.

Whenever robustness is incorporated into a model, an important concern is the loss in performance due to the optimization with respect to worst-cases. However, robust solutions immunized against significant drifts in parameters are obtainable in many cases, while incurring minuscule declines in performance compared to the deterministic optimal. This being so, robustness does not always have to imply loss in performance, be it small or large. Particularly, in the context of risk minimization, immunization complements the optimization perspective, rather than opposing. An elaborate modeling of parameter uncertainty can help alleviate the over-conservatism in the risk perspective and allow for a more moderate risk measure setting.

In the seminal work by Markowitz [2], variance of the portfolio return was used as the measure of portfolio risk. Variance as a risk measure has a limitation: it penalizes gains, i.e., upward deviations from the expected return, as well as the losses, i.e., downward deviations. Introduction of semivariance, i.e., the portion of variance due to downfalls below expected return, instead, alleviates this limitation of return variance as a risk measure. In this study we use lower partial moment (LPM) risk measures of different degrees:

$$LPM_m(X) = \mathbb{E}[(r - X)_+^m],$$

where r is the target/benchmark level for expected return, and m is the degree of the lower partial moment. For $m = 0$, LPM represents the probability of the return falling below the benchmark level r. For $m = 1$, LPM becomes the expected shortfall of the return below r, and for $m = 2$, LPM is the analogue of semivariance where calculations are based on deviations below the benchmark r rather than the return mean. Compared to semivariance, LPM has a strength in adjustability of investor preference in two axes: where semivariance regards as risk the downfall below expected return, LPM allows for a user set benchmark rate r, below which counts as a loss. With this handle, the ambition of the investor regarding return targets can be reflected by the risk measure. Secondly, with the degree of the LPM, m, the risk measure can be adjusted according to the risk aversion of the investor. What many portfolio selection models seek to address by a combination of weighted return and risk terms, can be implemented by a single LPM term in the objective function. Since LPM contains elements representing both the return target and the risk of loss, the portfolio selection models in this study that take (worst-case) LPMs as risk measures have merely one constraint: the equality balancing total allocation to assets with investor's wealth.

Despite lacking mathematically and computationally desirable properties, the *Value-at-Risk* (VaR) is standardized as a popular risk measure in finance applications. Similar to the manner losses are the focus of attention rather than gains, often losses at the extreme, i.e., those in the upper tail of the loss distribution are more important from a risk perspective. VaR measures risk at the upper extremity of the loss distribution, and is defined as follows:

$$VaR_\alpha(L) = \inf\{l \in \mathbb{R} : P(\{L > l\}) \leq 1 - \alpha\}.$$

In words, $VaR_\alpha(L)$ is the (minimum) loss value above which remains a probability chunk of (no more than) $1-\alpha$. Typical values for α are 0.95 and 0.99, thus under the tail beyond $VaR_\alpha(L)$ lies a probability of 0.05 or 0.01, respectively. Thus $VaR_\alpha(L)$ represents an extreme loss value, above which a small probability of loss realizations remain. However, VaR does not convey any information on the magnitude of losses that can be incurred beyond this threshold value. Especially when the loss distribution has a heavy upper tail, or with discrete distributions, the loss value represented by VaR can be a misleading underestimation. In addition, VaR lacks important mathematical properties sought in a coherent risk measure (as in [4]), such as subadditivity and convexity. Lacking subadditivity, VaR can penalize diversification, that is, even when investment is divided between two assets of independent return distributions, it is possible that VaR for this portfolio is larger than VaR for an investment in one of the assets only. Lack of convexity implies possibility of the existence of multiple minima and maxima, which brings significant computational complexity and is unpleasant from an optimization point of view.

Conditional Value-at-Risk, $CVaR$, although is a risk measure whose basic derivation is based on VaR, gathers in itself many desirable properties theoretically and computationally, is a coherent risk measure (satisfying monotonicity, subadditivity, homogeneity, concavity, and translational equivariance) and gives an account of the magnitude of the loss that can be incurred in the upper tail beyond VaR. $CVaR$ is defined as follows:

$$\begin{aligned} CVaR_\alpha(L) &= \frac{1}{1-\alpha}\mathbb{E}[L\mathbb{1}_{(VaR_\alpha(L),\infty)}] \\ &= \frac{1}{1-\alpha}\int\limits_{\{L>VaR_\alpha(L)\}} x dG_L(x), \end{aligned}$$

which is the expected loss conditioned on the loss being larger than corresponding VaR of the given confidence level α. Here, G_L is the loss distribution function. $CVaR$ admits an alternative form and can be computed by the minimization of the auxiliary function $F_L(\gamma) = \gamma + \frac{1}{1-\alpha}\mathbb{E}[(L-\gamma)_+]$:

$$\begin{aligned} CVaR_\alpha(L) &= \min_\gamma F_L(\gamma) \\ &= \min_\gamma \left(\gamma + \frac{1}{1-\alpha}\mathbb{E}[(L-\gamma)_+]\right) \\ &= \min_\gamma \left(\gamma + \frac{1}{1-\alpha}\int\limits_{\{L>\gamma\}} x dG_L(x)\right), \end{aligned}$$

which allows for linear/convex formulations of mathematical programs involving $CVaR$. VaR is the cause of intractability in many models, and substituting $CVaR$ instead often renders such models practically solvable. In favorable cases such as when asset returns are normal or from ellipsoidal distribution families, using $CVaR$ instead of VaR as the risk measure elicits solutions similar/identical to those of the VaR and variance risk models. Hence, in these cases, exploiting $CVaR$ makes possible obtaining the (approximate) solutions of these problems computationally intractable otherwise.

Among (worst case) VaR and $CVaR$ minimizing portfolio models studied here, the ones analogous to the LPM models contain an additional constraint besides that for wealth allocation: a constraint limiting from below the expected return of the portfolio. This is necessary for establishing the analogy, since VaR and $CVaR$ are measures of risk at the extreme, lacking a term for target gains such as the benchmark term in LPM.

In Chapter 3, we consider the problem of optimal portfolio choice using the second lower partial moment risk measure (LPM_2), for a market consisting of n risky assets and a riskless asset. When the mean return vector and variance/covariance matrix of the risky assets are specified without specifying a return distribution, we derive distributionally robust portfolio rules. We then address ambiguity in the mean return vector as well, in addition to distribution ambiguity, and derive a closed-form portfolio rule when the uncertainty in the return vector is modeled via an ellipsoidal uncertainty set. Our result also indicates a choice criterion for the radius of ambiguity of the ellipsoid. Using the adjustable robustness paradigm we extend the single period results to multiple periods, and derive closed-form dynamic portfolio policies which mimic closely the single period policy.

In Chapter 4, we consider the problem of optimal portfolio choice using the first lower partial moment (LPM_1) risk measure, again, for a market consisting of n risky assets and a riskless asset. When the mean return vector and variance/covariance matrix of the risky assets are specified without fixing a return distribution, and the investor is averse to uncertainty in the mean return estimate as well, we derive closed-form robust portfolio rules modeling the uncertainty in the return vector via an ellipsoidal uncertainty set. Using the adjustable robustness paradigm we extend the single period results to multiple periods in discrete time, and derive closed-form dynamic portfolio policies which replicate, mutatis mutandis, the single period policy.

In Chapter 5, we consider the problem of optimal portfolio choice using the CVaR and VaR measures for a market consisting of n risky assets and a riskless asset with short positions allowed. When the distribution of returns of risky assets is unknown but the mean return vector and variance/covariance matrix of the risky assets are fixed, we derive the distributionally robust portfolio rules. Then, we address uncertainty in the mean return vector in addition to distribution ambiguity, and derive the optimal portfolio rules when the uncertainty in the return vector is modeled via an ellipsoidal uncertainty set. In the presence of a riskless asset, the robust CVaR and VaR measures, coupled with a minimum mean return constraint, yield simple, mean-variance efficient optimal portfolio rules. In a market without the riskless asset, we obtain a closed-form portfolio rule that generalizes earlier results, without a minimum mean return restriction.

In Chapter 6, we have a change of perspective regarding uncertainty. Instead of an uncertainty defined by freedom of distribution type given mean/covariance information, we assume knowledge of a nominal multivariate probability distribution for asset returns. The uncertainty set is defined as ball around this nominal distribution where the (probability) metric used for defining the ball is the Kantorovich distance (also called Wasserstein metric). In Pflug et al.[5], it is demonstrated that when uncertainty, i.e., the radius of the ball, increases, the worst-case risk minimizing portfolio converges to the uniform portfolio where each asset receives equal portions of the wealth. Rather than a single specific risk measure, the main result is proved for a generic risk functional representing a class of convex, version independent risk measures. Following the result in Pflug et al.[5], we investigate convergence of the risky investment to uniform portfolio when a riskless asset is available. We first prove that investment for the wealth remaining after the allocation to the riskless asset converges to uniformity with increasing risk, as before. Next, we show that as risk increases, the total amount allocated to risky assets diminishes, while attaining a uniform shape. The result in Pflug et al.[5] is significant as a mathematical justification for the well appreciated naive diversification strategies in portfolio selection. Since the key factor in the success of the uniform portfolio is the increasing uncertainty radius, by exploiting a riskless asset, we assess the willingness to invest in the risky asset under market conditions involving high uncertainty. Indeed, the diminishing allocation to risky assets indicates that in a risky environment where uniform portfolio is (nearly) optimal, the investor tends to avoid the risky assets and withholds his/her wealth from the risky market.

In Chapter 7, we present the computational studies on the results in Chapters 3, 4 and 5, analyzing the behavior and comparing the performances of the solutions therein, under various settings of uncertainty.

Chapter 2

Literature Review

Mulvey et al. [1] in their pioneering work emphasize the necessity for a proactive strategy to immunize solutions against the negative effects on solution feasibility and optimality caused by drifts in the values of parameters, which are assumed to be fixed in deterministic mathematical programs. Robust Optimization (RO) is a proactive strategy, as opposed to the reactive assessments made post-optimally in Sensitivity Analysis. They present RO as a general framework for modeling uncertainty, based on a linear programming (LP) formulation with two classes of variables and classes of constraints. The class of variables named as "design variables" are those whose values have to be decided upon before the realization of uncertain parameters. "Control variables" on the other hand are flexible decision variables: decisions on these can be adjusted after the realization of uncertain parameters. One set of constraints in the LP formulation involves only the design variables, and has deterministic parameters. A second set of constraints has parameter ambiguity, and involves both design and control variables. The parameter uncertainty model is scenario based, there is a finite set of scenarios representing possible realizations of uncertain parameters, each with a specific probability weight. The constraints have to be satisfied in each scenario by the design variables and scenario specific control variables, up to an error term modeled by auxiliary variables. They compare the framework introduced with Stochastic Linear Programming, emphasizing the superiority of RO due to model flexibility and computational tractability.

Research on RO flares after the studies by Ben-Tal et al. [6, 7, 8] and El Ghaoui et al. [9, 10] draw attention on RO as a comprehensive methodology for addressing real world problems involving uncertainty: with an appropriate selection of uncertainty set structures, the robust counterparts of many optimization problems can be formulated as mathematical programs that are known to be computationally tractable. This amounts to a diverse collection of robust counterparts under different uncertainty set structures. With efficient interior point methods for solving convex/nonlinear optimization problems [11, 12], practical solvability is achieved beyond theoretical tractability results. The combination of model flexibility and practical solvability makes RO a widely applicable methodology on real world problems.

When parameter uncertainty is introduced to an LP via linear inequality based uncertainty sets, the resulting Robust Counterpart (RC) can be formulated as an LP. When the uncertainty sets are defined by conic quadratic inequalities (as in the case of ellipsoidal uncertainty sets) or linear matrix inequalities, the RC can be formulated as a Conic Quadratic Problem (CQP) or a Semidefinite Programming Problem (SDP), respectively. Robust LPs with polyhedral uncertainty sets can be formulated as LPs. When norm-based uncertainty is introduced to LPs, the resulting RC can be formulated as convex programming problems with constraints based on the dual of the norm defining the uncertainty [13]. Thus, when the uncertainty defining norms are l_1 or l_∞, the RC can be formulated as an LP, whereas with l_2-norm based uncertainty sets, the RC can be formulated as a Second Order Cone Program. RCs of Quadratically Constrained Quadratic Programs can be formulated as SDPs when parameter uncertainty is modeled by a single ellipsoid, while more complex geometries for the uncertainty set, such as polyhedron or intersection of ellipsoids, result in NP-hard RCs. RCs of Convex Quadratically Constrained Problems and SDPs are generally NP-hard even under simplest uncertainty set structures. Solution of these RCs are based on approximations [7, 14, 15, 16, 10]. The interested reader is referred to Ben-Tal and Nemirovski [17] and Bertsimas et al.[18], two studies extensively reviewing the field of RO.

Optimal portfolio selection is a major theme of the field of mathematical finance, starting with the work of H. Markowitz [19], and with seminal contributions from Merton [20, 21] and others; see e.g., Steinbach [22] for an excellent survey where both single and multiple period portfolio selection is studied. A more recent research effort concerns the problem of sensitivity of the optimal portfolio rules to probabilistic model assumptions about asset returns [23, 24, 25] where portfolios robust to ambiguity in distribution of asset returns are investigated. It is also a commonly recognized problem that the portfolio rules tend to be sensitive to imprecision in the assumed mean returns which are usually estimated from historical data [26, 27, 28]. Aversion to the uncertainty in mean return for optimal portfolio choice was addressed using the relatively recent technique of robust optimization [6, 29] in e.g., [30, 31, 32, 33, 34, 35, 36]. For recent reviews of robust optimization and portfolio selection see e.g., [37, 38, 39, 40]. A seminal reference on multi-period portfolio selection is Mossin [41] while a more recent work in the context of mean-variance portfolio selection is Li and Ng[42].

Portfolio optimization in single and multiple periods, using different criteria such as mean-variance and utility functions has been studied extensively, see e.g., [43, 44, 45, 46, 47, 42, 19, 48, 20, 49, 50, 51]. In particular, Hakansson [47] treats correlations between time periods while Merton [48, 20] concentrate on continuous-time problems. These references usually consider a stochastic model for the uncertain elements (asset returns) and study the properties of an optimal portfolio policy. An important tool here is stochastic dynamic programming.

The philosophy of RO [6, 29] is to treat the uncertain parameters in an optimization problem by confining their values to some uncertainty set without defining a stochastic model, and find a solution that satisfies the constraints of the problem regardless of the realization of the uncertain parameters in the uncertainty set. It has been applied with success to single period portfolio optimization; see e.g., [24, 25, 31, 36]. The usual approach is to choose uncertainty sets that lead to tractable convex programming problems that are solved numerically. In this study, we instead find closed-form portfolio rules. In the case of multiple period portfolio problems, RO was extended to adjustable robust optimization (ARO), an approach that does not resort to dynamic programming, is more flexible than the classical RO for sequential problems, but may lead to more difficult optimization problem instances; see [8, 52]. A related approach, which is data-driven with probabilistic guarantees and scenario generation, is explored in e.g. [53].

The optimal portfolio choice problem using lower partial moments risk measures under distribution ambiguity was studied by Chen et al.[23] in the case of n risky assets. The authors assumed that the mean return vector μ and variance-covariance matrix Γ of risky assets are fixed, and compute portfolios that are distributionally robust in the sense that they minimize a worst-case lower partial moment risk measure over all distributions with fixed first and second moment information. They obtained closed-form distributionally robust optimal portfolio rules. In Chapter 3, we focus our attention to a single risk measure: the mean semi-deviation from a target rate, i.e., the LPM_2 risk measure, and first extend the results in Chen et al.[23] to the case where a riskless asset is also included in the asset universe. The inclusion of the riskless asset in the asset universe simplifies considerably the optimal choice formula in some cases as we shall see in Theorem 3.1.1. A criticism leveled against the distributionally robust portfolios in Chen et al.[23] is the sensitivity of these portfolios to uncertainties or estimation errors in the mean return data, a case that we refer to as *Mean Return Ambiguity*; see Delage and Ye [24]. To address this issue, we analyze the problem when the mean return is subject to ellipsoidal uncertainty in addition to distribution ambiguity and derive a closed-form portfolio rule. Since the majority of contributions in robust portfolio optimization aim at providing convex optimization formulations our explicit portfolio rule constitutes a worthy addition to the literature. Our result is valid for choices of the ellipsoidal uncertainty (ambiguity radius) parameter ϵ not exceeding the optimal Sharpe ratio attainable in the market. Furthermore, the difference between the optimal mean semi-deviation risk under distribution ambiguity only and the same measure under joint uncertainty in distribution and mean return may also impose an optimal choice of ϵ, an observation which we illustrate numerically. For other related studies on portfolio optimization with distributional robustness, the reader is referred to [25, 34, 54]. We also obtain optimal dynamic portfolio rules using the adjustable robust optimization paradigm [8, 52] for both cases of distribution ambiguity and expected return ambiguity combined with distribution ambiguity. In the dynamic multi-period setting it was assumed as in Chen et al. [23] that uncertainty across periods is independent. The resulting portfolio rules are myopic replica of the single period results.

In Chapter 4, we substitute expected downside deviation from a target rate instead of mean semi-deviation, i.e., the LPM_1 risk measure instead of LPM_2, and extend the work in Chen et al.[23], again, in two directions: first a riskless asset was incorporated into the portfolio universe, which simplified considerably the optimal portfolio rules (c.f. Theorem 3.1.1, where the optimal policy for the expected downside deviation is simply twice the optimal policy for squared semi-deviation measure), and second we assumed aversion to mean return ambiguity in addition to distribution ambiguity on the part of the investor and derived closed-form portfolio rules in both static and dynamic settings. We give explicit optimal portfolio rules, deriving results in both static and dynamic adjustable robustness settings [8, 52], respectively in Theorem 4.1.2 and Theorem 4.2.1. Again, uncertainty is assumed to be independent across periods. Interestingly, the factor of 2 that binds the optimal policies in the case of distribution ambiguity only continues to remain valid in the case of joint distribution and mean return ambiguity. In the case of expected downside deviation as in the case of squared semi-deviation, the dynamic optimal portfolio policy is a kind of myopic policy in the following sense: the investor acts as if solving a single period problem in each period, i.e., he/she adjusts the end of horizon target using the discount factor (the riskless rate) and unless the adjusted target is achieved by placing all current wealth into the riskless account, he/she adopts a single period optimal policy for this adjusted wealth target. It is rare that the dynamic adjustable robust portfolio problems result in explicit portfolio policies as in the present study. Thus, Chapter 4 makes a contribution to the repertoire of explicitly solvable robust dynamic portfolio planning under risk measures.

The Value-at-Risk (VaR) is widely used in the financial industry as a downside risk measure. Since VaR does not take into account the magnitude of potential losses, the Conditional Value-At-Risk (CVaR), defined as the mean losses in excess of VaR, was proposed as a remedy and results usually in convex (linear) portfolio optimization problems [55, 56]. The purpose of the present work is to give an explicit solution to the optimal portfolio choice problem by minimizing the Conditional Value-at-Risk and Value-at-Risk measures under distribution and mean return ambiguity when short positions are allowed. Distribution ambiguity is understood in the sense that no knowledge of the return distribution for risky assets is assumed while the mean and variance/covariance are assumed to be known. The optimal portfolio choice problem using the aforementioned risk measures under distribution ambiguity and allowing short positions was studied by Chen et al. [23] in the case of n risky assets, extending the work of Zhu and Fukushima [57] where the authors treat robust portfolio choice under distribution ambiguity. Chen et al. [23] assumed that the mean return vector μ and variance-covariance matrix Γ of risky assets are known, and compute portfolios that are robust in the sense that they minimize the worst-case CVaR risk measure over all distributions with fixed first and second moment information. They obtained closed-form robust optimal portfolio rules. The reader is referred to [25, 34, 54] for other related studies on portfolio optimization with distributional robustness, and to [58] for a computational study of scenario based CVaR in portfolio optimization. A recent reference work on portfolio optimization (using the mean-variance approach as well as semi-variance and utility functions) in both single and multi-period frameworks is Steinbach [22]. In particular, in [34] Natarajan et al. study expected utility models in portfolio optimization under distribution ambiguity using a piecewise-linear concave utility function. They obtain bounds on the worst-case expected utility, and compute optimal portfolios by solving conic programs. They also relate their bounds to convex risk measures by defining a worst-case Optimized-Certainty-Equivalent (OCE) risk measure. It is well-known that one of the two risk measures used in the present study, namely CVaR, can be obtained using the OCE approach for a class of utility functions; see [59]. Thus the results of Chapter 5 complement the previous work of Natarajan et al. [34] by providing closed-form optimal portfolio rules for worst-case CVaR (and worst-case VaR) under both distribution and mean return ambiguity. In Chapter 5, we first extend, in Section 5.1, the results of Chen et al. [23] to the case where a riskless asset is also included in the asset universe. The inclusion of the riskless asset in the asset universe leads to extreme positions in the portfolio, which implies that the robust CVaR and VaR measures as given in the present study have to be utilized with a minimum mean return constraint in the presence of a riskless asset in order to yield closed-form optimal portfolio rules. The distribution robust portfolios of Chen et al. [23] are criticized in [24] for their sensitivity to uncertainties or estimation errors in the mean return data, i.e., Mean Return Ambiguity; see also [26, 27, 60] for studies regarding sensitivity of optimal portfolios to estimation errors. To (partially) address this issue, we analyze in Section 5.2 the problem when the mean return is subject to ellipsoidal uncertainty [6, 29, 32, 61, 33] in addition to distribution ambiguity, and derive a closed-form portfolio rule. The ellipsoidal uncertainty is regulated by a parameter that can be interpreted as a measure of confidence in the mean return estimate. In the presence of the riskless asset, a robust optimal portfolio rule under distribution and mean return ambiguity is obtained if the quantile parameter of CVaR or VaR measures is above a threshold depending on the optimal Sharpe ratio of the market and the confidence regulating parameter, or no such optimal rule exists (the problem is infeasible). The key to obtain optimal portfolio rules in the presence of a riskless under distribution and mean return ambiguity asset is again to include a minimum mean return constraint to trace the efficient robust CVaR (or robust VaR) frontier [22]. The incremental impact of adding robustness against mean return ambiguity in addition to distribution ambiguity is to alter the optimal Sharpe ratio of the market viewed by the investor. The investor views a smaller optimal Sharpe ratio decremented by the parameter reflecting the confidence of the investor in the mean return estimate. In the case the riskless asset is not included in the portfolio problem, in Section 5.3 we derive in closed form the optimal portfolio choice robust against distribution and ellipsoidal mean return ambiguity without using a minimum mean return constraint, which generalizes the result of Chen et al. [23] stated in the case of distribution ambiguity only, i.e., full confidence in the mean return estimate.

In Chapter 6, we have a non-parametric model of distribution uncertainty, rather than uncertainty based on known moments, albeit with some error in the moment information. We assume that ambiguity in asset return distribution is defined by a ball around a center constituted by a known nominal distribution. We adopt this form of uncertainty model with a specific investment strategy in the focus of our studies: naive diversification, i.e., investing in all assets with equal shares of the wealth. In Benartzi and Thaler [62], it is pointed out that naive diversification is a common practice both as a general heuristic of choice and an investment strategy. Behavioral experiments indicate that when subjects are asked to choose multiple items from a list of possible selections simultaneously, they tend to diversify their decisions, i.e., they pick as diverse a group of items as possible. In contrast, when subjects are asked to pick a single item from the list each time in a series of experiments, choices are confined to small group of items in the list. Such behavioral experiments are extended to decisions on investment plans. Investment choices of employees presented with a fictive mix of assets for their retirement saving plans display ubiquitous utilization of the naive diversification strategy, where the ratio of the total amount invested in stocks/bonds is in strong correlation with the ratio of the number of stocks/bonds offered in the asset mix. These behavioral studies are supported by investor behavior elicited from archives of investment history: investment is diversified in the direction of growth in the mix of assets available, and the pattern of naive diversification is observed in the correlation of total allocation to certain asset types with the prevalence of the asset type in the mixture. Such behavior contradicts the choice of rational investors as characterized in the portfolio selection literature, following, for instance, mean-variance portfolio rules. In addition to contradicting theoretic results in portfolio theory, the psychological bias towards naive diversification heuristics can be considered irrational since it is difficult to imagine that a rational model fits the preferences of the diversity of people expressing this choice. However, DeMiguel et al.[63] point out that the anticipation of some form of uncertainty in the environment might be intuitively leading people to naive diversification. The authors study 14 models, mostly the mean-variance model and its extensions, evaluating these based on 7 real market data sets. The portfolio rules are compared to the $1/N$ naive diversification rule as a benchmark. The result is fascinating: none of the 14 models investigated consistently outperforms naive diversification in terms of out-of-sample Sharpe ratio, certainty-equivalent return and turnover (trading volume) criteria. It is observed that uncertainty in parameters due to estimation errors outweigh the theoretical gains promised by models, and the authors point out the need for an unrealistic amount of data for the models to perform better than the $1/N$ rule. The result is significant for eliciting from real market data a justification for the fact that choosing naive diversification against optimal portfolio rules in the literature is not necessarily an irrational behavior, but indeed is a strategy hard to outperform. Pflug et al.[5] go one step further and provide theoretical/mathematical justification for the optimality of $1/N$ investment under a specific setting of uncertainty. Given the (known) nominal distribution for asset returns, they define the asset return distribution uncertainty by the ball of measures around the nominal distribution based on the Kantorovich distance. It is proven that as the radius of the ball, i.e. the level of uncertainty, increases, the optimal investment vector approaches the naive diversification solution - the $1/N$ uniform investment on all assets. The choice of metric is important for this result that constitutes a solid justification for the optimality of the naive diversification strategy. Kantorovich distance is bounded from below by the square of Prokhorov metric [64], which metricizes weak convergence on any separable metric space. The reverse property holds only if the probability space has finite diameter; a lower bound for Prokhorov metric based on Kantorovich distance exists only if the state space is bounded. Therefore, Kantorovich distance has stronger convergence properties in the sense that convergence of probability measures under Kantorovich distance implies, but is not necessarily implied by, weak convergence. From the converse point of view, a ball defined based on Kantorovich distance can be outer-approximated by a ball defined based on Prokhorov metric. This perspective is more relevant to the result in [5]. The convergence of optimal investment to $1/N$ investment on a ball based on Kantorovich distance does not necessarily imply the result when uncertainty is defined based on Prokhorov metric balls. Indeed, robust portfolio models assuming non-parametric distribution ambiguity are studied for uncertainty represented by balls based on other metrics. In Calafiore [65], a nominal discrete distribution and an uncertainty ball based on Kullback-Leibler divergence models the return distribution ambiguity. Worst-case optimal portfolio over all distributions in the uncertainty ball is computed for mean-variance and mean-absolute deviation risk measures. In the former, an interior point barrier method in conjunction with an analytic center cutting plane technique is used, and in the latter case, a line search algorithm is incorporated to the solution procedure, additionally. In Erdoğan and Iyengar [66], asset return distribution is considered to lie within a ball based on Prokhorov metric, and an ambiguous chance constrained problem is defined: the set of constraints have to be satisfied with a probability greater than a fixed threshold, and this has to hold for every probability measure inside the ball defining distribution ambiguity. The problem is approximated by robust sampling of probability distributions, which results in problem formulations having the same complexity as the nominal problem with certain distribution. In Chapter 7, we adopt the framework in Pflug et al.[5], for our interest in naive diversification strategies, and add a riskless asset to the investment environment that otherwise lacks an explicit model for a riskless asset. We first extent the result in Pflug et al.[5] showing that optimal risky investment approaches naive diversification for both positive and negative allocations remaining for the risky assets, after fixing a certain allocation for the riskless asset. While having this convergence effect, we show that increasing uncertainty radius causes the investor to steer away from the market of risky assets. With this analysis we aim to shed light on the desirability of investing in environments where the naive diversification heuristic becomes a plausible strategy.

Chapter 3

Mean Semi-Deviation from a Target and Robust Portfolio Choice under Distribution and Mean Return Ambiguity

The purpose of this chapter is to give an explicit solution to the optimal portfolio choice problem by minimizing the lower partial moment risk measure of mean semi-deviation from a target return under distribution and mean return ambiguity using a robust optimization (RO) approach. The results of this chapter are published in Journal of Computational and Applied Mathematics [67].

The plan of the chapter is as follows. In Section 3.1 we derive the optimal portfolio rules under distributional ambiguity for two measures of risk in the presence of a riskless asset. We study the multiple period adjustable robust portfolio rules in Section 3.2. In Section 3.3, we derive the optimal portfolio rule for the mean squared semi-deviation from a target measure under distributional ambiguity and ellipsoidal mean return uncertainty. We also discuss the optimal choice of the uncertainty radius for the mean return. The multiple period extension is given in Section 3.4.

3.1 Minimizing First and Second Lower Partial Moments in the Presence of a Riskless Asset: Single Period

The lower partial moment risk measure LPM_m for $m = 0, 1, 2$ is defined as

$$\mathbb{E}\left[r - X\right]_+^m$$

for a random variable X and target r. In addition to the n risky assets with given mean return μ and variance-covariance matrix Γ, we assume that a riskless asset with return rate $R < r$ exists. If $R \geq r$, then the benchmark rate is attained without risk, i.e. the lower partial moment LPM_m is minimized taking value 0 by investing entirely into the riskless asset. Denote by y the variable for the riskless asset, to handle it separately, and by e the n-dimensional vector of ones, the LPM_m minimizing robust portfolio selection model under distribution ambiguity is:

$$RPR_m = \min_{x,y} \sup_{\xi \sim (\mu,\Gamma)} \mathbb{E}\left[r - x^T\xi - yR\right]_+^m \tag{3.1}$$

$$\text{s.t} \quad x^Te + y = 1. \tag{3.2}$$

We use the notation $\xi \sim (\mu, \Gamma)$ to mean that random vector ξ belongs to the set whose elements have mean μ and variance-covariance matrix Γ. Now, we provide the analytical solutions of the riskless asset counterpart of the problem for $m = 1, 2$ (expected shortfall and expected squared semi-deviation from a target, respectively)following a similar line to the proof of LPM_m solutions in Chen et al. [23]. The optimal portfolio choice for $m = 0$, which corresponds to minimizing the probability of falling short of the target, is uninteresting in the presence of a riskless asset in comparison to the case of risky assets only since the optimal portfolio displays an extreme behavior (the components vanish or go to infinity). Therefore, we exclude this case in the theorem below.

Theorem 3.1.1. *Suppose $\Gamma \succ 0$ and $R < r$. The optimal portfolio in (3.1)-(3.2) is obtained as follows.*

1. For the case $m = 1$ the optimal portfolio rule is

$$x^* = \frac{2\tilde{r}}{1+H}\Gamma^{-1}\tilde{\mu}.$$

2. For the case $m = 2$ the optimal portfolio rule is

$$x^* = \frac{\tilde{r}}{1+H}\Gamma^{-1}\tilde{\mu},$$

where $H = \tilde{\mu}^T\Gamma^{-1}\tilde{\mu}$, $\tilde{\mu} = \mu - Re$ and $\tilde{r} = r - R$.

Proof. The equality constraint (3.2) can be dropped letting $y = 1 - x^T e$:

$$RPR_m = \min_x \sup_{\xi\sim(\mu,\Gamma)} \mathbb{E}\left[(r-R) - x^T(\xi - eR)\right]_+^m.$$

One-to-one correspondence between the sets of distributions

$$D = \{\pi | \mathbb{E}_\pi[\xi] = \mu, Cov_\pi[\xi] = \Gamma \succ 0\}$$

and

$$\tilde{D} = \{\pi | \mathbb{E}_\pi[\xi] = \mu - eR, Cov_\pi[\xi] = \Gamma \succ 0\}$$

can be easily established. Hence, the model can be written as:

$$RPR_m = \min_x \sup_{\xi\sim(\tilde{\mu},\Gamma)} \mathbb{E}\left[\tilde{r} - x^T\xi\right]_+^m$$

where $\tilde{r} = r - R$ and $\tilde{\mu} = \mu - eR$. To be able to use the bounds derived for LPM_m, the equivalent single-variable optimization model should be noted:

$$RPR_m = \min_x \sup_{\zeta\sim(x^T\tilde{\mu}, x^T\Gamma x)} \mathbb{E}\left[\tilde{r} - \zeta\right]_+^m.$$

The equivalence of the single-variable and multi-variable optimization models is based on the one-to-one correspondence of the sets of distributions that ξ and ζ may assume (a proof of this fact can be found in [23]).

We define objective functions with respect to mean return and variance, using the tight bounds provided for LPM_m, $m = 1, 2$ in [23]:

$$f_1(s,t) := \sup_{X\sim(s,t^2)} E\left[(\tilde{r} - X)_+^1\right] = \frac{\tilde{r} - s + \sqrt{t^2 + (\tilde{r} - s)^2}}{2}, \quad (3.3)$$

$$f_2(s,t) := \sup_{X\sim(s,t^2)} E\left[(\tilde{r} - X)_+^2\right] = \left[(\tilde{r} - s)_+\right]^2 + t^2. \quad (3.4)$$

Then, ν being the optimal value, the problem becomes:

$$\nu\left(RPR_m\right) = \min_x \left\{ f_m\left(x^T\tilde{\mu}, \sqrt{x^T\Gamma x}\right)\right\} \tag{3.5}$$

$$= \min_{s\in\mathbb{R}} \min_x \left\{ f_m\left(s, \sqrt{x^T\Gamma x}\right) \mid x^T\tilde{\mu} = s\right\} \tag{3.6}$$

Noting that f_m is non-decreasing in variance (t^2) for $m = 1, 2$, the inner optimization in (3.6) is solved by minimizing the variance:

$$\begin{aligned} \min_x \quad & x^T\Gamma x \\ s.t \quad & x^T\tilde{\mu} = s. \end{aligned}$$

Thus, we obtain

$$\begin{aligned} 2\Gamma x - u\tilde{\mu} &= 0, \\ x^T\tilde{\mu} &= s, \\ u &\in \mathbb{R}. \end{aligned}$$

Hence we have the optimal solution for the inner optimization:

$$x = \frac{u\Gamma^{-1}\tilde{\mu}}{2},$$

$$\frac{u\tilde{\mu}^T\Gamma^{-1}\tilde{\mu}}{2} = s,$$

which gives

$$\begin{aligned} u &= \frac{2s}{\tilde{\mu}^T\Gamma^{-1}\tilde{\mu}} \\ x_s^* &= \frac{s}{\tilde{\mu}^T\Gamma^{-1}\tilde{\mu}}\Gamma^{-1}\tilde{\mu}. \end{aligned}$$

Having found the optimal value for x given a fixed value of s, we can now define the objective function as a function of s only:

$$\begin{aligned} \phi_m(s) &:= f_m\left(\tilde{\mu}^T x_s^*, \sqrt{x_s^{*T}\Gamma x_s^*}\right) \\ &= f_m\left(s, \sqrt{x_s^{*T}\Gamma x_s^*}\right). \end{aligned}$$

Following the notation $H = \tilde{\mu}^T\Gamma^{-1}\tilde{\mu}$, we have:

$$\begin{aligned} x_s^{*T}\Gamma x_s^* &= \frac{s^2}{H^2}\tilde{\mu}^T\Gamma^{-1}\Gamma\Gamma^{-1}\tilde{\mu} \\ &= \frac{s^2}{H^2}\tilde{\mu}^T\Gamma^{-1}\tilde{\mu} \\ &= \frac{s^2}{H}, \end{aligned}$$

thus we write:

$$\begin{aligned}\phi_m(s) &= f_m\left(s, \sqrt{\frac{s^2}{H}}\right), \\ \nu(RPR_m) &= \min_s \{\phi_m(s)\}.\end{aligned}$$

Now we can seek s that minimizes $\phi_m(s)$, for cases $m = 1$, and 2 separately. For $m = 1$ we have:

$$\phi_1(s) = \frac{\tilde{r} - s + \sqrt{\frac{s^2}{H} + (\tilde{r} - s)^2}}{2}.$$

We are minimizing $\gamma_1(s) = \sqrt{\frac{s^2}{H} + (\tilde{r} - s)^2} - s$. The first order condition gives:

$$\gamma_1{}'(s) = \frac{\frac{s}{H} - \tilde{r} + s}{\sqrt{\frac{s^2}{H} + (\tilde{r} - s)^2}} - 1 = 0,$$

equivalent to

$$\left(\frac{s}{H} - \tilde{r} + s\right)^2 = \frac{s^2}{H} + (\tilde{r} - s)^2 \text{ and} \tag{3.7}$$

$$\frac{s}{H} - \tilde{r} + s \geq 0. \tag{3.8}$$

Equation (3.7) has two roots, one of which is 0, not satisfying (3.8). The other root, $s = \frac{2\tilde{r}}{1 + \frac{1}{H}}$, satisfies (3.8) and is the minimizer of $\gamma_1(s)$, since $\gamma_1'(s)$ is negative to the left and positive to the right of this value. To see this, we let $a := 1 + \frac{1}{H}$, and write:

$$\begin{aligned}\gamma_1'(s) &= \frac{\frac{s}{H} - \tilde{r} + s}{\sqrt{\frac{s^2}{H} + (\tilde{r} - s)^2}} - 1 \\ &= \frac{s\left(1 + \frac{1}{H}\right) - \tilde{r}}{\sqrt{s^2\left(1 + \frac{1}{H}\right) + \tilde{r}^2 - 2\tilde{r}s}} - 1 \\ &= \frac{as - \tilde{r}}{\sqrt{as^2 + \tilde{r}^2 - 2\tilde{r}s}} - 1\,,\end{aligned} \tag{3.9}$$

observing ${\gamma_1}'(s) \leq -1$ if $s \leq \dfrac{\tilde{r}}{a}$. If $s \in \left(\dfrac{\tilde{r}}{a}, \dfrac{2\tilde{r}}{a}\right)$, then the nominator in (3.9) is positive, but:

$$\begin{aligned}
(as - \tilde{r})^2 &= a^2 s^2 - 2as\tilde{r} + \tilde{r}^2 \\
&= as(as - 2\tilde{r}) + \tilde{r}^2 \\
&< s(as - 2\tilde{r}) + \tilde{r}^2 \\
&= as^2 - 2\tilde{r}s + \tilde{r}^2 \\
&= \left(\sqrt{as^2 - 2\tilde{r}s + \tilde{r}^2}\right)^2 .
\end{aligned} \tag{3.10}$$

We have $as - \tilde{r} < \sqrt{as^2 + \tilde{r}^2 - 2\tilde{r}s}$, thus ${\gamma_1}'(s) < 0$. In inequality (3.10), note that $\tilde{r} > 0$, $a > 1$, $s > 0$ and $as - 2\tilde{r} < 2\tilde{r} - 2\tilde{r} = 0$. If $s > \dfrac{2\tilde{r}}{a}$, then $as - 2\tilde{r} > 0$; and inequality (3.10) is in the opposite direction. It follows that ${\gamma_1}'(s) > 0$ if $s > \dfrac{2\tilde{r}}{a}$, and ${\gamma_1}'(s) < 0$ if $s < \dfrac{2\tilde{r}}{a}$, hence $s_1^* = \dfrac{2\tilde{r}}{a}$ is the unique minimizer of $\gamma_1(s)$ and $\phi_1(s)$.

Finally, for $m = 2$, the minimizer of $\phi_2(s)$ is $s_2^* = \dfrac{\tilde{r}}{1 + \dfrac{1}{H}}$. $\phi_2(s)$ can be defined in piecewise form:

$$\begin{aligned}
\phi_2(s) &= \left[(\tilde{r} - s)_+\right]^2 + \frac{s^2}{H} \\
&= \begin{cases} (\tilde{r} - s)^2 + \dfrac{s^2}{H} & \text{if } s < \tilde{r} \\ \dfrac{s^2}{H} & \text{if } s \geq \tilde{r}, \end{cases}
\end{aligned}$$

and has continuous first derivative:

$$\phi_2'(s) = \begin{cases} 2s\left(1 + \dfrac{1}{H}\right) - 2\tilde{r} & \text{if } s < \tilde{r} \\ \dfrac{2s}{H} & \text{if } s \geq \tilde{r}. \end{cases}$$

$\phi_2'(s)$ is positive if $s \geq \tilde{r}$, and $2s\left(1 + \dfrac{1}{H}\right) - 2\tilde{r}$ is an affine function of s with positive slope that takes value 0 at $s_2^* = \dfrac{\tilde{r}}{1 + \dfrac{1}{H}} < \tilde{r}$. Therefore s_2^* is the unique minimizer of $\phi_2(s)$, with negative first derivative to the left and positive to the right side. $\square$

The constant H that appears in the optimal portfolio rules is the highest attainable Sharpe ratio in the market, see e.g. [28]. This constant plays an important role in Theorem 3.3.1 in Section 3.4.

Comparing our results for $m = 1, 2$ to the corresponding result (Theorem 2.5) of [23] we notice that the optimal portfolio rules look much simpler. In fact, the optimal portfolio rule in case $m = 1$ is exactly twice the optimal portfolio rule in case $m = 2$. This simple relationship between the two rules can be attributed to the fact that the case $m = 2$ is more conservative in that it punishes more severely the deviations from target compared to the case $m = 1$. As the two optimal portfolios are almost identical up to a constant multiplicative factor, and it is easier to deal with the case $m = 2$ we shall concentrate on that case in the next section.

An immediate but slight generalization is to allow a budget W_0 instead of 1 in (3.2). This has the effect of redefining $\tilde{r}$ as $r - W_0 R$.

3.2 Multi-period Portfolio Rule under Distribution Ambiguity with a Riskless Asset

In the present section we shall extend the result of the previous section for the case $m = 2$ to a multiple period adjustable robustness setting. The reason we limit ourselves to $m = 2$ is the fact that we shall deal exclusively with that case in the rest of the chapter when we consider ambiguity in mean return.

Consider now, for the sake of illustration, a multiple period problem with three periods, i.e., $T = 3$. The situation is the following. At the beginning of time period $t = 1$, the investor has a capital W_0 which he allocates among n risky assets with mean return vector μ_1 and variance/covariance matrix Γ_1 and riskless rate R (for the sake of simplicity, assumed constant throughout the entire horizon) according to the expected semi-deviation from a target risk measure. His endowment is W_1 at the beginning of period $t = 2$ where he faces expected return vector μ_2, variance/covariance matrix Γ_2 where he allocates his wealth again to obtain at the end of period $t = 2$ a wealth W_2. This wealth is again invested into risky assets with expected return vector μ_3 and matrix Γ_3. It is assumed that all matrices Γ_i $i = 1, 2, 3$ are invertible. It is also assumed that random variables in each period are independent from those in other periods, and in this setting a learning model for moment information through periods is not incorporated.

Let the portfolio positions be represented by vectors $x_t \in \mathbb{R}^n$ for $t = 1, 2, 3$ (risky assets), and by scalars y_t, for $t = 1, 2, 3$ (riskless asset). For a chosen end-of-horizon target wealth r, the adjustable robust portfolio selection problem is defined recursively as follows:

$$V_3 = \min_{x_3, y_3} \max_{\xi_3 \sim (\mu_3, \Gamma_3)} \mathbb{E}[r - \xi_3^T x_3 - R y_3]_+^2$$

subject to

$$e^T x_3 + y_3 = W_2$$

$$V_2 = \min_{x_2, y_2} \max_{\xi_2 \sim (\mu_2, \Gamma_2)} \mathbb{E}[V_3]$$

subject to

$$e^T x_2 + y_2 = W_1$$

$$V_1 = \min_{x_1, y_1} \max_{\xi_1 \sim (\mu_1, \Gamma_1)} \mathbb{E}[V_2]$$

subject to

$$e^T x_1 + y_1 = W_0.$$

The idea is that while for an observer at the beginning of period 1, the wealths W_1 and W_2 are random quantities, the realized wealth $\tilde{W}_1$ say, is a known quantity at the beginning of period 2. The same is true of realized wealth, $\tilde{W}_2$ say, at the beginning of period 3. These observations allows to adjust the portfolio according to realized random information instead of selecting all portfolios for all periods at the very beginning.

We begin solving the problem above from period $t = 3$. Using Theorem 3.1.1, we have that

$$x_3^* = \left(\frac{r - W_2 R}{1 + H_3} \right) \Gamma_3^{-1} \tilde{\mu}_3$$

where $\tilde{\mu}_3 = \mu_3 - Re$ and $H_3 = \tilde{\mu}_3^T \Gamma_3^{-1} \tilde{\mu}_3$. We substitute this quantity into the objective function and obtain the expression

$$V_3 = \frac{1}{(1+H_3)^2}[(r - W_2 R)_+^2 + (r - W_2 R)^2 H_3].$$

Now we need to find the supremum of the expectation of V_3 over all random variables $\xi_2 \sim (\mu_2, \Gamma_2)$, i.e., we need to solve the problem

$$\sup_{\xi_2 \sim (\mu_2, \Gamma_2)} \frac{1}{(1+H_3)^2} \mathbb{E}[(r - R^2 W_1 - R(\xi_2 - Re)^T x_2)_+^2 + H_3 (r - R^2 W_1 - R(\xi_2 - Re)^T x_2)^2]$$

after substituting for y_2. This maximization problem is solved using a simple extension of Lemma 1 of [23] (its proof is verbatim repetition of the proof of Lemma 1 of [23], thus omitted):

Lemma 3.2.1. *Let the random variable X have mean and variance (μ, σ^2). Then we have for any $\alpha, \beta \in \mathbb{R}$*

$$\sup_{X \sim (\mu, \sigma^2)} \mathbb{E}[\alpha(r - X)_+^2 + \beta(r - X)^2] = (\alpha + \beta)\sigma^2 + \beta(r - \mu)^2 + \alpha(r - \mu)_+^2.$$

Applying the above result gives the function

$$\frac{1}{(1+H_3)^2}\left[R^2 x_2^T \Gamma_2 x_2 + H_3(r - R^2 W_1 - R(\mu_2 - Re)^T x_2)^2 \right.$$
$$\left. + (r - R^2 W_1 - R(\mu_2 - Re)^T x_2)_+^2\right]$$

to be minimized over x_2 using the techniques in the proof of Theorem 3.1.1. This results in the solution

$$x_2^* = \left(\frac{r - W_1 R^2}{R(1 + H_2)}\right) \Gamma_2^{-1} \tilde{\mu}_2,$$

where $\tilde{\mu}_2 = \mu_2 - Re$ and $H_2 = \tilde{\mu}_2^T \Gamma_2^{-1} \tilde{\mu}_2$. Repeating the above steps for V_2 (the details are left as an exercise) we obtain the solution x_1^* as

$$x_1^* = \left(\frac{r - W_0 R^3}{R^2(1 + H_1)}\right) \Gamma_1^{-1} \tilde{\mu}_1,$$

with $\tilde{\mu}_1 = \mu_1 - Re$ and $H_1 = \tilde{\mu}_1^T \Gamma_1^{-1} \tilde{\mu}_1$. The above process can be routinely generalized to arbitrary integer T time periods. Thus we have the following theorem.

Theorem 3.2.1. *Let $r - W_{t-1}R^{T-t+1} > 0$ for $t = 1, \ldots, T$. The adjustable robust multi-period portfolio rule using the expected squared semi-deviation from a target wealth r risk measure in a T periods setting is*

$$x_t^* = \left(\frac{r - W_{t-1}R^{T-t+1}}{R^{T-t}(1 + H_t)}\right)\Gamma_t^{-1}\tilde{\mu}_t,$$

for $t = 1, 2, \ldots, T$ where $\tilde{\mu}_t = \mu_t - Re$ and $H_t = \tilde{\mu}_t^T\Gamma_t^{-1}\tilde{\mu}_t$.

Compared to Theorem 3.1 of [23] our result is so much simpler, and gives a myopic dynamic portfolio policy in the following sense. The single period optimal portfolio policy consists in setting a target excess wealth beyond that which could be obtained by putting all the present wealth in the riskless asset: $r - W_0R$. Dividing this excess target wealth by the optimal Sharpe ratio H plus one, one obtains the coefficient in the optimal rule. A similar formula is given in the previous theorem for the multi-period case. Note that each term $\frac{r - W_{t-1}R^{T-t+1}}{R^{T-t}}$ has the following economic meaning: the investor looks at the end of the current period t and sets the excess wealth target equal to

$$\frac{r}{R^{T-t}} - W_{t-1}R$$

which is exactly the discounted target wealth value at time $t+1$ minus the wealth that would be obtained if the current wealth W_t was kept in the riskless account for one period. If this number is equal to zero or is negative, then the final target can simply be achieved by investing the current wealth into the riskless asset for the rest of the horizon; hence the optimal position in risky assets would be zero for the remaining periods. If this excess target remains positive for all periods t, divided by the optimal period t Sharpe ratio H_t plus one we have the optimal rule for each period. In other words, it is as if the investor is solving at each time period the following problem

$$\min_{x_t, y_t} \max_{\xi_t \sim (\mu_t, \Gamma_t)} \mathbb{E}[\frac{r}{R^{T-t}} - \xi_t^T x_t - Ry_t]_+^2$$

subject to

$$e^T x_t + y_t = W_{t-1}.$$

3.3 Distribution and Expected Return Ambiguity: Single Period

It is well-documented that the optimal portfolios may be quite sensitive to inaccuracies in the mean return vector; see e.g. [26, 27, 24]. To address this issue we consider now the problem

$$RPRR_2 = \min_{x,y} \sup_{\xi\sim(\bar{\mu},\Gamma),\, \bar{\mu}\in U_\mu} \mathbb{E}\left[r - x^T\xi - yR\right]_+^2 \tag{3.11}$$

$$\text{s.t} \quad x^Te + y = 1. \tag{3.12}$$

where we define the ellipsoidal uncertainty set $U_{\bar{\mu}} = \{\bar{\mu} | \|\Gamma^{-1/2}(\bar{\mu} - \mu_{\text{nom}})\|_2 \leq \sqrt{\epsilon}\}$ for the mean return denoted $\bar{\mu}$, where μ_{nom} denotes a nominal mean return vector which can be taken as the available estimate of mean return. The ellipsoidal representation of uncertain parameters is now a well-established choice in the robust optimization literature, and in particular in portfolio optimization; see e.g. [6, 29, 32, 61, 33] for discussion and motivations for the choice of an ellipsoidal set. We refer to the positive parameter ϵ as the *radius of ambiguity.*

Theorem 3.3.1. *Let $\tilde{r} > 0$ and $\mu^* = \mu_{nom} - Re$. Then the optimal portfolio rule x^* to (3.11)-(3.12) is given by*

$$x^* = \frac{\tilde{r}(\sqrt{H} + \sqrt{\epsilon})(H - \epsilon)}{\sqrt{H}[(\sqrt{H} + \sqrt{\epsilon})^2 + (H - \epsilon)^2]}\Gamma^{-1}\mu^*.$$

provided that $\epsilon < H$. If $\epsilon \geq H$ all wealth is invested into the riskless asset.

Proof. Using Remark 2.8 of [23] we can pose the problem (3.11)-(3.12) as

$$\min_{x,s} (\tilde{r} - s)_+^2 + x^T\Gamma x$$

subject to

$$s \leq \min_{\bar{\mu}\in U_{\bar{\mu}}} x^T(\bar{\mu} - Re)$$

after elimination of the variable y as in the proof of Theorem 3.1.1. Writing the constraint explicitly we get the problem

$$\min_{x,s} (\tilde{r} - s)_+^2 + x^T\Gamma x$$

subject to the conic constraint

$$s \leq x^T \mu^* - \sqrt{\epsilon}\sqrt{x^T \Gamma x},$$

where $\mu^* = \mu_{\text{nom}} - Re$. Assuming $\tilde{r} > s$ and the constraint to be active, the stationarity equations of the KKT necessary and sufficient optimality conditions give (note that Slater's condition holds trivially)

$$s = \frac{\lambda}{2} + \tilde{r}$$

where λ is a non-negative Lagrange multiplier, and

$$2\Gamma x - \lambda(\mu^* - \frac{\sqrt{\epsilon}}{\sqrt{x^T \Gamma x}} \Gamma x) = 0,$$

under the hypothesis that $x \neq 0$. Defining $\sigma = \sqrt{x^T \Gamma x}$ we rewrite the last equation as

$$2\Gamma x - \lambda(\mu^* - \frac{\sqrt{\epsilon}}{\sigma} \Gamma x) = 0,$$

which yields

$$x^* = \frac{\lambda \sigma}{2\sigma + \lambda\sqrt{\epsilon}} \Gamma^{-1} \mu^*.$$

We have two equations that allow to solve for λ and σ: the first equation comes from the definition of σ and gives

$$\frac{\lambda^2 \sigma^2}{(2\sigma + \lambda\sqrt{\epsilon})^2} H = \sigma^2,$$

and the second equation that comes from the conic constraint

$$\tilde{r} - \frac{\lambda}{2} = \frac{\lambda \sigma}{2\sigma + \lambda\sqrt{\epsilon}} H - \sqrt{\epsilon} \frac{\lambda \sigma}{2\sigma + \lambda\sqrt{\epsilon}} \sqrt{H}.$$

We solve the first equation for λ holding σ fixed, and obtain the roots

$$\frac{2(\sqrt{\epsilon} + \sqrt{H})\sigma}{H - \epsilon}, \frac{2(\sqrt{\epsilon} - \sqrt{H})\sigma}{H - \epsilon}.$$

The second root is always negative while the first root is positive for $\epsilon < H$ provided σ is positive, which we assume to be the case. Solving for σ from the second equation above, we obtain after some straightforward simplification

$$\sigma = \frac{\tilde{r}(H - \epsilon)(\sqrt{H} + \sqrt{\epsilon})}{(\sqrt{H} + \sqrt{\epsilon})^2 + (H - \epsilon)^2}$$

which is positive provided $\epsilon < H$. Substituting back into the expression for x^* and simplifying we get the desired expression. If $H \leq \epsilon$ then our hypothesis that $\sigma > 0$ is false, hence the KKT conditions do not yield an optimal solution, except when $H = \epsilon$ in which case the optimal choice is $x^* = 0$.

On the other hand, assuming $\tilde{r} \leq s$, we obtain $\lambda = 0$ and $x^* = 0$ which gives $s \leq 0$. But this is a contradiction since $\tilde{r} > 0$, so this case is impossible. □

The optimal portfolio is a mean-variance efficient portfolio. The optimal Sharpe ratio H serves as an upper bound for the radius of ambiguity. We note that when $\epsilon = 0$ we recover exactly the optimal portfolio rule in case 2 of Theorem 3.1.1 with $\mu_{\text{nom}} = \mu$ where μ was defined in Section 3.2. We define for ease of notation

$$\kappa(\epsilon) = \frac{(\sqrt{H} + \sqrt{\epsilon})(H - \epsilon)}{\sqrt{H}[(\sqrt{H} + \sqrt{\epsilon})^2 + (H - \epsilon)^2]},$$

which is the critical factor introduced by robustness against ambiguity in the mean return vector. This quantity $\kappa(\epsilon)$ is a decreasing function in ϵ as illustrated in Figure 3.3.

Consider now the difference $RPR_2 - RPRR_2$ in the risk measures. The respective values are given as

$$RPR_2 = \left[(r - R)(1 - \frac{H}{1 + H})\right]_+^2 - (r - R)^2 \frac{H}{1 + H},$$

and

$$RPRR_2 = \left[(r - R)(1 - \kappa(\epsilon)H - \kappa(\epsilon)H\sqrt{\epsilon})\right]_+^2 - (r - R)^2 \kappa(\epsilon)H.$$

We expect the difference $RPR_2 - RPRR_2$ to be positive at least for a range of values of ϵ. Since the function $RPRR_2$ is complicated to analyze, we provide a numerical example with $H = 0.24$, $r = 1.05$ and $R = 1.03$ in Figure 3.3. The gain in mean squared semi-deviation risk reaches a peak for some value ϵ^* of ϵ and then starts to fall. This behavior could guide the choice for an appropriate value of the radius of ambiguity ϵ. We also note that as H increases the maximizer ϵ^* shifts to the right as well. This can be seen by comparing with Figure 3.3 where we used $H = 0.54$ all other parameters being equal.

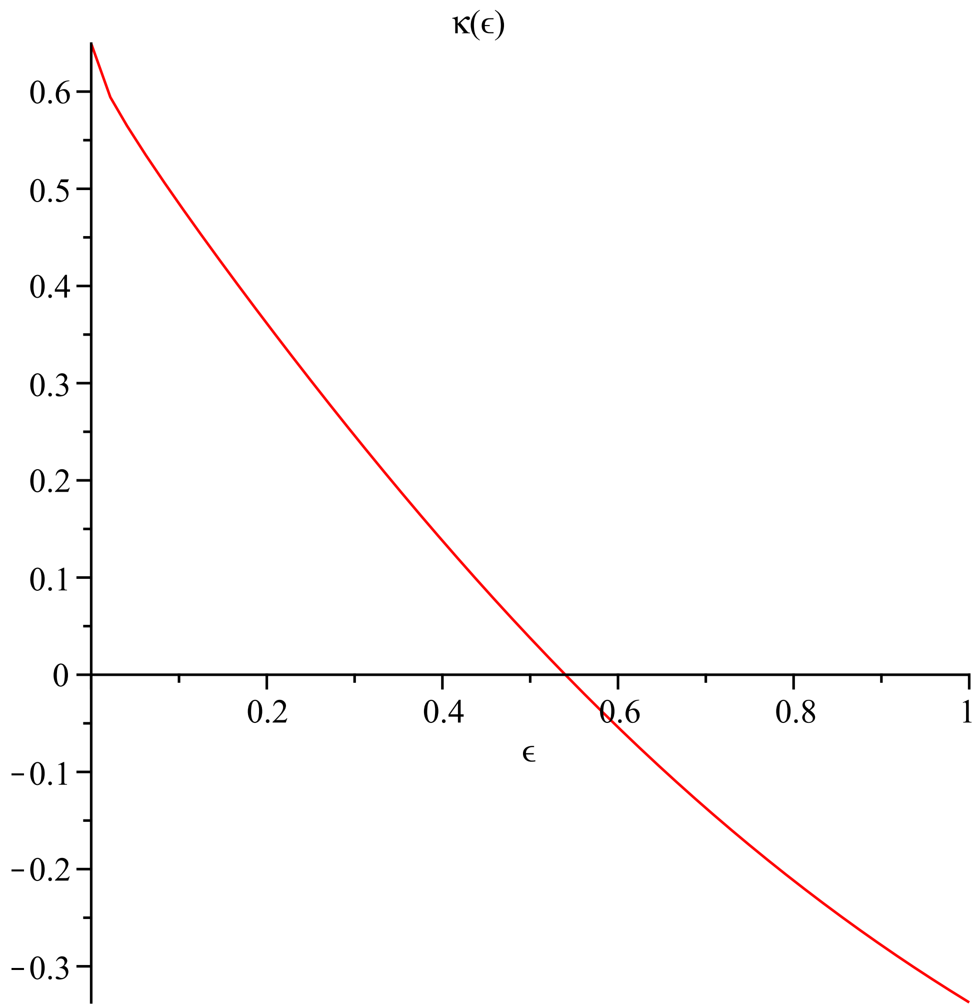

Figure 3.1: $\kappa(\epsilon)$ as a function of the ellipsoidal uncertainty radius ϵ with $H = 0.54$.

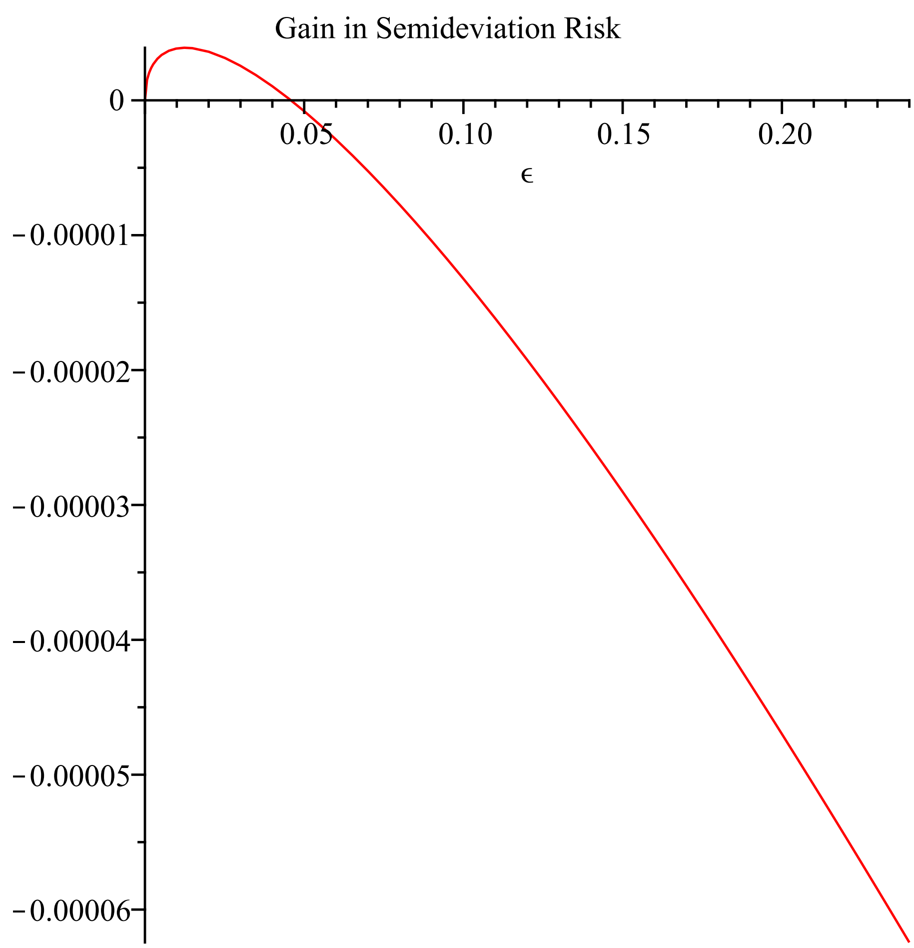

Figure 3.2: Gain in mean semi-deviation risk as a function of the ellipsoidal uncertainty radius ϵ with $H = 0.24$, $r = 1.05$, $R = 1.03$.

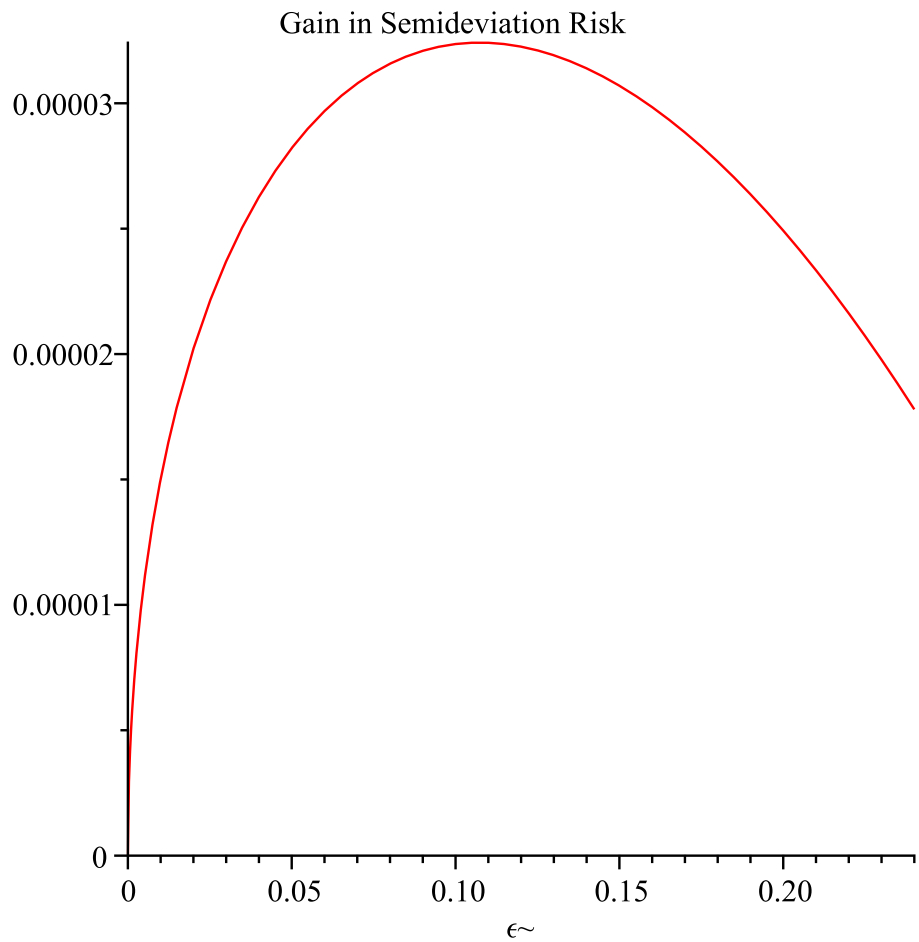

Figure 3.3: Gain in mean semi-deviation risk as a function of the ellipsoidal uncertainty radius ϵ with $H = 0.54$, $r = 1.05$, $R = 1.03$.

3.4 Multi-period Case under Distribution and Mean Return Ambiguity with a Riskless Asset

In the present section we shall extend the result of the previous section to a multiple period adjustable robustness setting.

Consider again a multiple period problem with three periods, i.e., $T = 3$. The situation is the following. At the beginning of time period $t = 1$, the investor has a capital W_0 which he allocates among n risky assets with mean return vector $\bar{\mu}_1$ and variance/covariance matrix Γ_1 and riskless rate R (for the sake of simplicity, assumed constant throughout the entire horizon) according to the expected semi-deviation from a target risk measure. His wealth is W_1 at the beginning of period $t = 2$ where he faces expected return vector $\bar{\mu}_2$, variance/covariance matrix Γ_2 where he allocates his wealth again to obtain at the end of period $t = 2$ a wealth W_2. This wealth is again invested into risky assets with expected return vector $\bar{\mu}_3$ and matrix Γ_3. It is assumed that all matrices Γ_i $i = 1, 2, 3$ are positive definite. The vectors $\bar{\mu}_1$, $\bar{\mu}_2$ and $\bar{\mu}_3$ are assumed uncertain, and to belong to respective ellipsoids

$$U^t_{\bar{\mu}} = \{\bar{\mu} | \|\Gamma^{-1/2}(\bar{\mu} - \mu^t_{\text{nom}})\|_2 \leq \sqrt{\epsilon_t}\}$$

for the mean return denoted $\bar{\mu}$, where μ^t_{nom} denotes a nominal mean return vector and ϵ_t are positive constants for $t = 1, 2, 3$. Let the portfolio positions be represented by vectors $x_t \in \mathbb{R}^n$ for $t = 1, 2, 3$ (risky assets), and by scalars y_t, for $t = 1, 2, 3$ (riskless asset). For a chosen end-of-horizon target wealth r, the adjustable robust portfolio selection problem is defined recursively as follows:

$$V_3 = \min_{x_3, y_3} \max_{\xi_3 \sim (\bar{\mu}_3, \Gamma_3)} \max_{\bar{\mu}_3 \in U^3_{\bar{\mu}}} \mathbb{E}[r - \xi_3^T x_3 - R y_3]^2_+$$

subject to

$$e^T x_3 + y_3 = W_2$$

$$V_2 = \min_{x_2, y_2} \max_{\xi_2 \sim (\bar{\mu}_2, \Gamma_2)} \max_{\bar{\mu}_2 \in U^2_{\bar{\mu}}} \mathbb{E}[V_3]$$

subject to

$$e^T x_2 + y_2 = W_1$$

$$V_1 = \min_{x_1, y_1} \max_{\xi_1 \sim (\bar{\mu}_1, \Gamma_1)} \max_{\bar{\mu}_1 \in U^1_{\bar{\mu}}} \mathbb{E}[V_2]$$

subject to

$$e^T x_1 + y_1 = W_0.$$

The computations are tedious but similar to those in Section 3.3 with the exception that one has to use the proof technique of Theorem 3.3.1 in solving the period minimization problems. By way of illustration we look at the periods $t = 3$ and $t = 2$. For $t = 3$, we have immediately the optimal portfolio rule from Theorem 3.3.1:

$$x_3^* = \frac{(r - W_2 R)(\sqrt{H_3} + \sqrt{\epsilon_3})(H_3 - \epsilon_3)}{\sqrt{H_3}[(\sqrt{H_3} + \sqrt{\epsilon_3})^2 + (H_3 - \epsilon_3)^2]} \Gamma_3^{-1} \mu_3^*$$

where $\mu_3^* = \mu^3_{\text{nom}} - Re$ provided that $\epsilon_3 < H_3$ and $r > W_2 R$. Substituting this solution back to the objective function and using Lemma 3.2.1 of Section 3.3 to evaluate the sup over $\xi_2 \sim (\bar{\mu}_2, \Gamma_2)$ we obtain the function

$$(A\tilde{r}_2 - AR(\bar{\mu}_2 - Re)^T x_2)_+^2 + A^2 R^2 x_2^T \Gamma_2 x_2 + B^2 R^2 x_2^T \Gamma_2 x_2 + B^2 (\tilde{r}_2 - R(\bar{\mu}_2 - Re)^T x_2)^2,$$

where $\tilde{r}_2 = r - W_1 R^2$, $A = 1 - \kappa_3 H_3 + \kappa_3 \sqrt{H_3 \epsilon_3}$ $B = \kappa_3 \sqrt{H_3}$, and

$$\kappa_3 = \frac{(\sqrt{H_3} + \sqrt{\epsilon_3})(H_3 - \epsilon_3)}{\sqrt{H_3}[(\sqrt{H_3} + \sqrt{\epsilon_3})^2 + (H_3 - \epsilon_3)^2]}.$$

Now we evaluate the sup of the above expression over $\bar{\mu}_2 \in U^2_{\bar{\mu}}$ and obtain the optimization problem

$$\min_{x_2, s} (A\tilde{r}_2 - ARs)_+^2 + A^2 R^2 x_2^T \Gamma_2 x_2 + B^2 R^2 x_2^T \Gamma_2 x_2 + B^2 (\tilde{r}_2 - Rs)^2$$

subject to

$$s \leq \mu_2^{*T} x_2 - \sqrt{\epsilon_2} \sqrt{x_2^T \Gamma_2 x_2}$$

where $\mu_2^* = \mu^2_{\text{nom}} - Re$. We solve this problem using the KKT conditions exactly as in the proof of Theorem 3.3.1 and obtain

$$x_2^* = \frac{(r - W_1 R^2)(\sqrt{H_2} + \sqrt{\epsilon_2})(H_2 - \epsilon_2)}{R\sqrt{H_2}[(\sqrt{H_2} + \sqrt{\epsilon_2})^2 + (H_2 - \epsilon_2)^2]} \Gamma_1^{-1} \mu_2^*,$$

under the condition $\epsilon_2 < H_2$ and $r > W_1 R^2$. Repeating the above steps one more time for $t = 1$ we arrive at

$$x_1^* = \frac{(r - W_0 R^3)(\sqrt{H_1} + \sqrt{\epsilon_1})(H_1 - \epsilon_1)}{R^2 \sqrt{H_1}[(\sqrt{H_1} + \sqrt{\epsilon_1})^2 + (H_1 - \epsilon_1)^2]} \Gamma_1^{-1} \mu_1^*$$

under the conditions $\epsilon_1 < H_1$ and $r > W_0 R^3$. Hence, generalizing the previous derivation to an arbitrary number of periods we obtain the following theorem.

Theorem 3.4.1. *Let $r - W_{t-1} R^{T-t+1} > 0$ and $\mu_t^* = \mu_{nom}^t - Re$ and $H_t = (\mu_t^*)^T \Gamma_t^{-1} \mu_t^*$ for $t = 1, \ldots, T$. Then the adjustable robust multi-period optimal portfolio rule under distribution and mean return ambiguity in a T periods setting is given by*

$$x_t^* = \frac{(r - W_{t-1} R^{T-t+1})(\sqrt{H_t} + \sqrt{\epsilon_t})(H_t - \epsilon_t)}{R^{T-t} \sqrt{H_t}[(\sqrt{H_t} + \sqrt{\epsilon_t})^2 + (H_t - \epsilon_t)^2]} \Gamma_t^{-1} \mu_t^*,$$

provided that $\epsilon_t < H_t$ for $t = 1, \ldots, T$.

Notice that for $\epsilon_t = 0$ and $\mu_{\text{nom}}^t = \mu_t$ as defined in Section 3.3, we obtain the dynamic portfolio rule of Theorem 3.2.1. Remarks similar to those made after Theorem 3.2.1 in Section 3.3 also hold for the dynamic portfolio rule of Theorem 3.4.1, i.e., the excess target wealth is chosen exactly as described at the end of Section 3.3. The remaining part of the portfolio rule is identical to the single period rule. In other words, it is as if the investor is solving at each time period the following problem

$$\min_{x_t, y_t} \max_{\xi_t \sim (\bar{\mu}_t, \Gamma_t), \bar{\mu}_t \in U_{\bar{\mu}}^t} \mathbb{E}[\frac{r}{R^{T-t}} - \xi_t^T x_t - R y_t]_+^2$$

subject to

$$e^T x_t + y_t = W_{t-1}.$$

The conditions $r - W_{t-1} R^{T-t+1} > 0$ also make economic sense because if at any time point the condition fails to hold it means that we have achieved a wealth figure that can equal or exceed the target wealth r by staying in the riskless asset for the remaining portion of the horizon until end of period T.

3.5 Conclusion

In this chapter we derived explicit optimal portfolio rules in single period and multiple period investment environments using the risk measure of expected squared semi-deviation from a target under both distribution ambiguity of asset returns and uncertainty in the mean return vector. We incorporated a riskless asset into the asset universe, which considerably simplifies the portfolio rules. In multiple periods, the optimal portfolio rule is a myopic replica of the single period rule in the following sense. If the target has not been reached, it is as if the investor is solving at every period a single period problem with some adjustments to the target. The case of expected shortfall (i.e., $m = 1$) under distribution and mean return ambiguity will be addressed in the next chapter.

Chapter 4

Expected Downside Deviation and Robust Portfolio Choice under Distribution and Mean Return Ambiguity

This chapter is organized as follows. In Section 4.1 we recall the problem setting and derive the static optimal portfolio rule for the risk measure of expected downside deviation from a target when the investor is averse to both distribution and mean return ambiguity which is represented using an ellipsoidal uncertainty set. In Section 4.2, we examine the multiperiod version of the robust portfolio selection problem under expected downside deviation criterion in the face of distribution and mean return ambiguity.

4.1 Minimizing Absolute Semi-deviation from a Target in the Presence of a Riskless Asset under Distribution and Mean Return Ambiguity: Single Period

The lower partial moment risk measure LPM_m for $m = 0, 1, 2$ is defined as

$$\mathbb{E}\left[r - X\right]_+^m$$

for a random variable X and target r. In addition to the n risky assets with given mean return μ and variance-covariance matrix Γ, we assume that a riskless asset with return rate $R < r$ is available. If $R \geq r$, then the benchmark rate is attained without risk, i.e. the lower partial moment LPM_m is minimized taking value 0 by investing entirely into the riskless asset. We use y to denote the variable for the holdings of the riskless asset, and by e the n-dimensional vector of ones, the LPM_m minimizing robust portfolio selection model under distribution ambiguity is:

$$RPR_m = \min_{x,y} \sup_{\xi \sim (\mu,\Gamma)} \mathbb{E}\left[r - x^T\xi - yR\right]_+^m \tag{4.1}$$

$$\text{s.t} \quad x^T e + y = 1, \tag{4.2}$$

where variables x and y represent the amount of investment into the risky assets and the riskless asset respectively; and ξ is the vector of return rates for the risky assets. We use the notation $\xi \sim (\mu, \Gamma)$ to mean that random vector ξ belongs to the set whose elements have mean μ and variance-covariance matrix Γ. In Chapter 3, closed-form solutions of the above problem for $m = 0, 1, 2$ were investigated. Let us recall the result of Lemma 2.2 of [23] which will be useful for us as well (it is credited to Scarf [68]).

Lemma 4.1.1. *[[23], Lemma 2.2]*

$$\sup_{X \sim (\mu,\sigma^2)} \mathbb{E}\left[r - X\right]_+ = \frac{r - \mu + \sqrt{\sigma^2 + (r-\mu)^2}}{2}.$$

While the lemma concerns univariate random variables, the equivalence of the univariate and multivariate cases is based on the one-to-one correspondence of the sets of distributions that ξ and X may assume, a result established in Lemma 2.4 of [23]. Now, we recall from Chapter 3 the analytical solution of the riskless asset counterpart of the problem for $m = 1$ (expected downside deviation) which is the focus of the present chapter.

Theorem 4.1.1 (Chapter 3, Theorem 3.1.1)**.** *Suppose $\Gamma \succ 0$ and $R < r$. The optimal portfolio in (4.1)-(4.2) for $m = 1$ is obtained as follows:*

$$x^* = \frac{2\tilde{r}}{1+H^2}\Gamma^{-1}\tilde{\mu}.$$

where $H^2 = \tilde{\mu}^T\Gamma^{-1}\tilde{\mu}$, $\tilde{\mu} = \mu - Re$ and $\tilde{r} = r - R$.

The constant H^2 that appears in the optimal portfolio rules is the highest attainable Sharpe ratio in the market, see e.g. [28]. This constant plays an important role in the sequel.

An immediate but slight generalization is to allow a budget W_0 instead of 1 in (4.2). This has the effect of redefining $\tilde{r}$ as $r - W_0 R$.

For the case $m = 2$, the optimal portfolio rule turns out to be half of the optimal rule for $m = 1$, namely

$$x^* = \frac{\tilde{r}}{1+H^2}\Gamma^{-1}\tilde{\mu},$$

c.f. Theorem 3.1.1 in Chapter 3. We shall see below that the proportionality factor of 2 among the two cases $m = 1$ and $m = 2$ is preserved even under distribution and ellipsoidal mean return ambiguity.

We consider now the problem

$$RPRR_1 = \min_{x,y} \sup_{\xi\sim(\bar{\mu},\Gamma),\, \bar{\mu}\in U_{\bar{\mu}}} \mathbb{E}\left[r - x^T\xi - yR\right]_+ \tag{4.3}$$

$$\text{s.t} \quad x^T e + y = 1. \tag{4.4}$$

where we define the ellipsoidal uncertainty set $U_{\bar{\mu}} = \{\bar{\mu} | \|\Gamma^{-1/2}(\bar{\mu} - \mu_{\text{nom}})\|_2 \leq \epsilon\}$ for the mean return denoted $\bar{\mu}$, where μ_{nom} denotes a nominal mean return vector taken e.g., as the available estimate of mean return. The ellipsoidal representation of uncertain parameters is now a well-justified and established practice in the robust optimization literature, and in particular in portfolio optimization; see e.g. [6, 29, 31, 32, 33] for discussion, results and motivations for the choice of an ellipsoidal set. We refer to the positive parameter ϵ as the *radius of uncertainty.*

Theorem 4.1.2. *Let $\tilde{r} > 0$ and $\mu^* = \mu_{nom} - Re$. Then the optimal portfolio rule x^* to (4.3)-(4.4), i.e., the expected downside deviation minimizing portfolio rule under distribution and ellipsoidal mean return ambiguity is given by*

$$x^* = \frac{2\tilde{r}(H-\epsilon)}{H + H(H-\epsilon)^2}\Gamma^{-1}\mu^*, \tag{4.5}$$

provided that $\epsilon < H$. If $\epsilon \geq H$ all wealth is invested into the riskless asset.

Proof. To prove the result we first recall from [23, 67] that the problem under distribution ambiguity only is posed as

$$RPR_1 = \min_x \frac{1}{2}\left(\tilde{r} - x^T(\bar{\mu} - Re) + \sqrt{x^T\Gamma x + (\tilde{r} - x^T(\bar{\mu} - Re))^2}\right).$$

Now, we examine the problem of interest in the form posed in [23]:

$$RPRR_1 = \min_x \sup_{\bar{\mu}\in U_{\bar{\mu}}} \frac{1}{2}\left(\tilde{r} - s + \sqrt{x^T\Gamma x + (\tilde{r} - s)^2}\right),$$

subject to

$$s \leq x^T(\bar{\mu} - Re) - \sqrt{x^T\Gamma x}$$

which is transformed, using the usual duality transformation of robust optimization (see e.g., [6, 29]) into the equivalent problem (referred to as (RP1) for later reference):

$$RPRR_1 = \min_{x,s} \frac{1}{2}\left(\tilde{r} - s + \sqrt{x^T\Gamma x + (\tilde{r} - s)^2}\right),$$

subject to

$$s \leq x^T\mu^* - \epsilon\sqrt{x^T\Gamma x}$$

which is a convex second-order cone optimization problem in variables x, s, where $\mu^* = \mu_{\text{nom}} - Re$. To solve the problem we make the guess that the optimal solution of the above problem is a mean-variance efficient portfolio rule, i.e., we posit a solution of the form

$$x = k\Gamma^{-1}\mu^*,$$

for some constant k. Hence, we have $x^T\mu^* = kH^2$ where $H^2 = (\mu^*)^T\,\Gamma^{-1}\mu^*$ and $x^T\Gamma x = k^2H^2$. By substituting the above quantities into the function and assuming that the constraint is tight at optimality we eliminate the variable s and the constraint, and we reduce the problem into one in the scalar variable k:

$$\min_k f(k) \equiv \frac{1}{2}\left[\tilde{r} - kH^2 + \epsilon kH + \sqrt{k^2 + (\tilde{r} - kH^2 + \epsilon kH)^2}\right]$$

where we assumed $k > 0$, and hence $\sqrt{x^T\Gamma x} = kH$. Under the supposition that we have a non-zero quantity under the square root, we take the derivative of f:

$$f' = -\frac{1}{2}H^2 + \frac{1}{2}\epsilon H + \frac{1}{2}\frac{kH^2 + (\tilde{r} - kH^2 + \epsilon kH)(-H^2 + \epsilon H)}{\sqrt{k^2 + (\tilde{r} - kH^2 + \epsilon kH)^2}}.$$

Let us define the quantity under the square root to be $\mathfrak{s}$:

$$\mathfrak{s} \equiv \sqrt{k^2 + (\tilde{r} - kH^2 + \epsilon kH)^2}.$$

Hence, from the first-order condition $f' = 0$ we obtain, after some algebra, solving for k:

$$k = \frac{(\mathfrak{s} + \tilde{r})(H - \epsilon)}{H + (H^2 - \epsilon H)(H - \epsilon)}. \tag{4.6}$$

Substituting this expression into the equation

$$\mathfrak{s}^2 \equiv k^2 + (\tilde{r} - kH^2 + \epsilon kH)^2$$

after some evident algebraic manipulation we obtain a quadratic equation in $\mathfrak{s}$ with the two roots equal to $\pm\tilde{r}$. Since the only allowable root is the positive root $\tilde{r} > 0$ ($\mathfrak{s}$ was defined to be the square root of some number), we substitute the value $\tilde{r}$ into expression (4.6) and obtain

$$k = \frac{2\tilde{r}(H - \epsilon)}{H + H(H - \epsilon)^2}.$$

We had begun under the hypothesis that $k > 0$, which is fulfilled if $H > \epsilon$. The condition $H > \epsilon$ also ensures that the quantity under square root in f, namely $k^2 + (\tilde{r} - kH^2 + \epsilon kH)^2$ is positive at the candidate solution. Now, we can verify that the hypothesized solution

$$x^* = \frac{2\tilde{r}(H - \epsilon)}{H + H(H - \epsilon)^2}\Gamma^{-1}\mu^*$$

is indeed a global minimizer of the second-order cone program (RP1). The first-order optimality conditions (FOC) of (RP1) are obtained from the Lagrange function:

$$L(x, s, \lambda) = \frac{1}{2}\left(\tilde{r} - s + \sqrt{x^T\Gamma x + (\tilde{r} - s)^2}\right) + \lambda(s - x^T\mu^* + \epsilon\sqrt{x^T\Gamma x}).$$

Using the hypothesis that the constraint is binding we have immediately

$$s^* = \frac{2\tilde{r}(H - \epsilon)^2}{1 + (H - \epsilon)^2}.$$

The FOC with respect to s give the identity

$$\lambda^* = \frac{1}{2} + \frac{1}{2}\frac{\tilde{r} - s^*}{\sqrt{(x^*)^T\Gamma x^* + (\tilde{r} - s^*)^2}},$$

which results in the expression

$$\lambda^* = \frac{1}{2} + \frac{1}{2}\frac{1 - (H - \epsilon)^2}{\sqrt{H^2 - 2H\epsilon + \epsilon^2 + 1}}.$$

Since $H^2 - 2H\epsilon + \epsilon^2 + 1$ is equal to $(1 - (H - \epsilon)^2)^2 + 4(H - \epsilon)^2$, we have that $\lambda^* > 0$. It remains to verify that the FOC with respect to x are verified at the point x^*, s^*, λ^*, i.e., the set of equations

$$\frac{1}{2}\frac{\Gamma x}{\sqrt{x^T\Gamma x + (\tilde{r} - s)^2}} - \lambda\mu^* + \lambda\epsilon\frac{\Gamma x}{\sqrt{x^T\Gamma x}} = 0,$$

holds true at the triplet (x^*, s^*, λ^*). Substituting the expressions for x^*, s^*, λ^* one obtains that the left hand side is equal to

$$\frac{1}{2}(\frac{H - \epsilon - H + \epsilon}{H})\mu^*$$

which is obviously equal to zero (this follows since the equation $H^2 - 2H\epsilon + \epsilon^2 + 1 = 0$ in ϵ has no real root, and hence the polynomial is always positive, and H is positive by positive definiteness of Γ).

Now, we examine the case $k < 0$. In this case, we have the function of k to be minimized:

$$\min_k f(k) \equiv \frac{1}{2}\left[\tilde{r} - kH^2 - \epsilon kH + \sqrt{k^2 + (\tilde{r} - kH^2 - \epsilon kH)^2}\right].$$

Carrying out the analysis as above (it is impossible that the quantity under the square root is zero under the hypothesis that $k < 0$), we obtain the solution parametric in $\mathfrak{s}$:

$$k = \frac{(H^2 + \epsilon H)(\tilde{r} + \mathfrak{s})}{H^2 + (H^2 + \epsilon H)^2}.$$

Solving for $\mathfrak{s}$ from the quadratic equation:

$$\mathfrak{s}^2 = \frac{(H^2 + \epsilon H)^2(\tilde{r} + \mathfrak{s})^2}{[H^2 + (H^2 + \epsilon H)^2]^2} H^2,$$

we obtain the two roots equal to $\pm\tilde{r}$ where the acceptable positive root is discarded because it gives a positive k, which is a contradiction. Hence, a negative value for k is impossible. Therefore, the only possibility is to take a zero position in risky assets and invest all wealth into the riskless asset. □

In case the initial capital is different from one, but equal to W_0, say, we redefine $\tilde{r}$ as $r - W_0 R$.

We immediately note that if the radius of ambiguity is zero, i.e., for $\epsilon = 0$ in Theorem 4.1.2, the optimal portfolio rule collapses to the optimal rule given in Theorem 4.1.1 for the case of robustness against distribution ambiguity only.

Compare the result, i.e., the optimal portfolio (4.5) of Theorem 4.1.2 to the corresponding result, c.f. Theorem 3 of [67], for $m = 2$ where we obtained the optimal portfolio rule under robustness to both distribution and mean return ambiguity:

$$x = \frac{\tilde{r}(H - \epsilon)}{H[1 + (H - \epsilon)^2]} \Gamma^{-1} \mu^*.$$

The two respective portfolio rules differ only by a multiplicative constant factor of 2. Furthermore, they are both valid under the condition $H > \epsilon$.

4.2 Multiple Periods and Adjustable Robust Portfolios

Now we shall extend the result of the previous section to a multiple period setting using the concept of adjustable robust solutions; see [8, 52] for primary references on adjustable robust solutions.

For ease of exposition (the case for a general integer number of periods is easy to obtain repeating the analysis given below) we consider a multiple period portfolio selection problem with three periods, i.e., $T = 3$ using the expected downside deviation criterion and under the assumption of aversion to distribution and mean return ambiguity. The situation is the following. At the beginning of time period $t = 1$, the investor has a capital W_0 which he allocates among n risky assets with mean return vector $\bar{\mu}_1$ and variance/covariance matrix Γ_1 and riskless rate R (for the sake of simplicity, assumed constant throughout the entire horizon) according to the expected downside deviation from a target risk measure. His wealth is W_1 at the beginning of period $t = 2$ where he faces expected return vector $\bar{\mu}_2$, variance/covariance matrix Γ_2 where he allocates his wealth again to obtain at the end of period $t = 2$ a wealth W_2. This wealth is again invested into risky assets with expected return vector $\bar{\mu}_3$ and matrix Γ_3. It is assumed that all matrices Γ_i $i = 1, 2, 3$ are positive definite. The vectors $\bar{\mu}_1$, $\bar{\mu}_2$ and $\bar{\mu}_3$ are assumed uncertain, and to belong to respective ellipsoids

$$U^t_{\bar{\mu}} = \{\bar{\mu} | \|\Gamma^{-1/2}(\bar{\mu} - \mu_{t,\text{nom}})\|_2 \leq \epsilon_t\}$$

for the mean return denoted $\bar{\mu}$, where $\mu_{t,\text{nom}}$ denotes a nominal mean return vector for period t and ϵ_t are positive constants for $t = 1, 2, 3$. Let the portfolio positions be represented by vectors $x_t \in \mathbb{R}^n$ for $t = 1, 2, 3$ (risky assets), and by scalars y_t, for $t = 1, 2, 3$ (riskless asset). For simplicity we assume a constant riskless return R throughout the horizon although the extension to a varying rate is straightforward. For a chosen end-of-horizon target wealth r, the adjustable robust portfolio selection problem is defined recursively as follows:

$$V_3 = \min_{x_3, y_3} \max_{\xi_3 \sim (\bar{\mu}_3, \Gamma_3),\ \bar{\mu}_3 \in U^3_{\bar{\mu}}} \mathbb{E}[r - \xi_3^T x_3 - R y_3]_+$$

subject to

$$e^T x_3 + y_3 = W_2$$

$$V_2 = \min_{x_2, y_2} \max_{\xi_2 \sim (\bar{\mu}_2, \Gamma_2),\, \bar{\mu}_2 \in U_{\bar{\mu}}^2} \mathbb{E}[V_3]$$

subject to

$$e^T x_2 + y_2 = W_1$$

$$V_1 = \min_{x_1, y_1} \max_{\xi_1 \sim (\bar{\mu}_1, \Gamma_1),\, \bar{\mu}_1 \in U_{\bar{\mu}}^1} \mathbb{E}[V_2]$$

subject to

$$e^T x_1 + y_1 = W_0.$$

The *adjustable robust* nature of the solution stems from the observation that the portfolio decisions of the investor are not made *a priori* at the beginning of the planning horizon for all periods. Instead, a robust portfolio selection is made for one period, and after the realized wealth is observed, an new robust selection is made, and so on until the end of the horizon. In stochastic optimization, the aforementioned adaptability (hence, the term adjustable) of the decision sequence to stochastic information that becomes available as time progresses, is known as *non-anticipativity.*

Let us begin the solution process backwards by solving

$$V_3 = \min_{x_3, y_3} \max_{\xi_3 \sim (\bar{\mu}_3, \Gamma_3),\, \bar{\mu}_3 \in U_{\bar{\mu}}^3} \mathbb{E}[r - \xi_3^T x_3 - R y_3]_+$$

subject to

$$e^T x_3 + y_3 = W_2$$

where for an investor at the beginning of period 3, W_2 is no longer stochastic, it has already been revealed. The solution according to Theorem 4.1.2 is

$$x_3^* = \frac{2\tilde{r}_3(H_3 - \epsilon_3)}{H_3 + H_3(H_3 - \epsilon_3)^2} \Gamma_3^{-1} \mu_3^*, \tag{4.7}$$

where $H_3 = \sqrt{\mu^{*T}_3 \Gamma_3^{-1} \mu_3^*}$, $\mu_3^* = \mu_{3,\text{nom}} - Re$, $\tilde{r}_3 = r - W_2 R$, provided that $\epsilon_3 < H_3$ and $r > W_2 R$, which we assume to be the case. For ease of notation, we define $K_3 = \frac{2(H_3 - \epsilon_3)}{H_3 + H_3(H_3 - \epsilon_3)^2}$. Hence, after some straightforward calculations we obtain

$$V_3 = \frac{1}{2}\tilde{r}_3 \left(1 - K_3 H_3^2 + \epsilon_3 K_3 H_3 + \sqrt{K_3^2 H_3^2 + (1 - K_3 H_3^2 + \epsilon_3 K_3 H_3)^2} \right).$$

Using a further encoding $\alpha_3 = 1 - K_3 H_3^2 + \epsilon_3 K_3 H_3$ and $\beta_3 = K_3 H_3$ we have

$$V_3 = \frac{1}{2}\left(\alpha_3 + \sqrt{\beta_3^2 + \alpha_3^2}\right)\tilde{r}_3.$$

Notice that due to our assumption that $\tilde{r}_3 = r - W_2 R > 0$, there is no loss of information if we re-write V_3 as

$$V_3 = \frac{1}{2}\left(\alpha_3 + \sqrt{\beta_3^2 + \alpha_3^2}\right)\max(0, \tilde{r}_3).$$

since this is precisely the form needed to invoke Lemma 4.1.1. On the other hand if $\tilde{r}_3 = r - W_2 R \leq 0$ then the end target wealth is already attained, and the investor would simply put all the wealth into the riskless asset. Now, we are ready to pass to period 2. i.e.,

$$V_2 = \min_{x_2, y_2} \max_{\xi_2 \sim (\bar{\mu}_2, \Gamma_2),\ \bar{\mu}_2 \in U_{\bar{\mu}}^2} \mathbb{E}[V_3]$$

subject to

$$e^T x_2 + y_2 = W_1.$$

For the investor at the beginning of period 2, the quantity W_1 is known. After the usual simplifications, invoking Lemma 4.1.1 (Lemma 2.2 of [23]), applying the ellipsoidal robustness transformation and eliminating the conic constraint, we have that

$$V_2 = \min_{x_2} \frac{c_3}{2}\left[\tilde{r}_2 - R x_2^T \mu_2^* + R\epsilon\sqrt{x_2^T \Gamma_2 x_2} \right.$$
$$\left. + \sqrt{R^2 x_2^T \Gamma_2 x_2 + (\tilde{r}_2 - R x_2^T \mu_2^* + R\epsilon\sqrt{x_2^T \Gamma_2 x_2})^2}\right],$$

where we defined the constant $c_3 \equiv \frac{1}{2}(\alpha_3 + \sqrt{\beta_3^2 + \alpha_3^2})$ to unburden the formula, and $\tilde{r}_2 = r - W_1 R^2$. Now, the solution of the above problem is carried out exactly as in Theorem 4.1.2, i.e., by guessing the form of the solution (i.e., $x_2 = k_2 \Gamma_2^{-1} \mu_2^*$) , and verifying the correctness of the guess afterwards (all the details of Theorem 4.1.2 reoccur here, so they are not repeated). The solution is obtained as

$$x_2^* = \frac{2\tilde{r}_2(H_2 - \epsilon_2)}{RH_2 + RH_2(H_2 - \epsilon_2)^2}\Gamma_2^{-1}\mu_2^*, \tag{4.8}$$

where $H_2 = \sqrt{\mu^{*T}_2 \Gamma_2^{-1} \mu_2^*}$, under the condition $H_2 > \epsilon_2$ and $\tilde{r}_2 > 0$. Substituting this solution into the objective function and defining $K_2 = \frac{2(H_2-\epsilon_2)}{RH_2+RH_2(H_2-\epsilon_2)^2}$ we obtain

$$V_2 = \frac{c_3}{2}\left(\alpha_2 + \sqrt{\beta_2^2+\alpha_2^2}\right)\tilde{r}_2$$

where $\alpha_2 = 1 - \frac{2H_2(H_2-\epsilon)}{1+(H_2-\epsilon_2)^2} + \frac{2\epsilon_2(H_2-\epsilon_2)}{1+(H_2-\epsilon_2)^2}$ and $\beta_2 = \frac{2(H_2-\epsilon_2)}{1+(H_2-\epsilon_2)^2}$. At this point we are ready to work on period 1:

$$V_1 = \min_{x_1,y_1} \max_{\xi_1\sim(\bar{\mu}_1,\Gamma_1),\, \bar{\mu}_1\in U_{\bar{\mu}}^1} \mathbb{E}[V_2]$$

subject to

$$e^T x_1 + y_1 = W_0.$$

Hence we are facing the problem

$$V_1 = \min_{x_1} \max_{\xi_1\sim(\bar{\mu}_1,\Gamma_1)\, \bar{\mu}_1\in U_{\bar{\mu}}^1} \mathbb{E}[\frac{c_2c_3}{2}\max(0,\tilde{r}_2)]$$

where we defined the constant $c_2 = \alpha_2 + \sqrt{\beta_2^2+\alpha_2^2}$ for convenience, and replaced without loss of generality $\tilde{r}_2$ by $\max(0,\tilde{r}_2)$ (recall that $\tilde{r}_2$ is positive by hypothesis since otherwise the end-of-horizon target is attained simply by placing the current wealth into the riskless asset). After eliminating the variable y_1, invoking Lemma 4.1.1 and applying the ellipsoidal robustness transformation and eliminating the conic constraint as in the proof of Theorem 4.1.2 one more time, we obtain the transformed problem

$$\begin{aligned} V_1 = \min_{x_1} \frac{c_2c_3}{4}\Bigg[&\tilde{r}_1 - R^2x_1^T\mu_1^* + \epsilon R^2\sqrt{x_1^T\Gamma_1x_1} \\ &+\sqrt{R^4x_1^T\Gamma_1x_1 + (\tilde{r}_1 - R^2x_1^T\mu_1^* + \epsilon R^2\sqrt{x_1^T\Gamma_1x_1})^2}\Bigg], \end{aligned}$$

where $\tilde{r}_1 = r - R^3W_0$. We again guess a solution of the form $x_1 = k_1\Gamma_1^{-1}\mu_1^*$ and verify the exactness as in the proof of Theorem 4.1.2. Since all the details are exactly as in the proof of Theorem 4.1.2 we give directly the optimal portfolio rule as

$$x_1^* = \frac{2\tilde{r}_1(H_1-\epsilon_1)}{R^2H_1 + R^2H_1(H_1-\epsilon_1)^2}\Gamma_1^{-1}\mu_1^*, \tag{4.9}$$

where $H_1 = \sqrt{\mu^{*T}_1 \Gamma_1^{-1} \mu_1^*}$, under the condition $H_1 > \epsilon_1$ and $\tilde{r}_1 > 0$. Generalizing the above process to a total of T periods, we can summarize the above result in the following.

Theorem 4.2.1. *Let $r - W_{t-1}R^{T-t+1} > 0$ and $\mu_t^* = \mu_{t,nom} - Re$ and $H_t = (\mu_t^*)^T \Gamma_t^{-1} \mu_t^*$ for $t = 1, \ldots, T$. Then, the adjustable robust multi-period optimal expected downside deviation portfolio rule under distribution and mean return ambiguity in a T periods setting is given by*

$$x_t^* = \frac{2(r - W_{t-1}R^{T-t+1})(H_t - \epsilon_t)}{R^{T-t}[H_t + H_t(H_t - \epsilon_t)^2]} \Gamma_t^{-1} \mu_t^*,$$

provided that $\epsilon_t < H_t$ for $t = 1, \ldots, T$. If at any time $t \in \{1, \ldots, T\}$ the investor has $r - W_{t-1}R^{T-t+1} \leq 0$ then it is optimal to invest all wealth W_{t-1} into the riskless asset until the end of the horizon.

It is noteworthy that for $\epsilon_t = 0$ (no aversion to mean return ambiguity) and taking $\mu_{t,\text{nom}} = \mu_t$ for each period t, we obtain the dynamic portfolio rule for the case of adjustable robustness in the face of distribution ambiguity only.

Corollary 4.2.1. *Let $r - W_{t-1}R^{T-t+1} > 0$ and $H_t^2 = (\mu_t^*)^T \Gamma_t^{-1} \mu_t^*$ for $t = 1, \ldots, T$. Then the adjustable robust multi-period optimal expected downside deviation portfolio rule under distribution ambiguity in a T periods setting is given by*

$$x_t^* = \frac{r - W_{t-1}R^{T-t+1}}{R^{T-t}(1 + H_t^2)} \Gamma_t^{-1} \mu_t,$$

for $t = 1, \ldots, T$. If at any time $t \in \{1, \ldots, T\}$ the investor has $r - W_{t-1}R^{T-t+1} \leq 0$ then it is optimal to invest all wealth W_{t-1} into the riskless asset until the end of the horizon.

Corollary 4.2.1 could also be obtained by starting the computations from scratch (but, without ellipsoidal uncertainty) in the context of adjustable robustness as we have done above in this section. e.g., in our three-period setting for a chosen end-of-horizon target wealth r, the adjustable robust portfolio selection problem is defined recursively as follows:

$$V_3 = \min_{x_3, y_3} \max_{\xi_3 \sim (\bar{\mu}_3, \Gamma_3)} \mathbb{E}[r - \xi_3^T x_3 - R y_3]_+$$

subject to

$$e^T x_3 + y_3 = W_2$$

$$V_2 = \min_{x_2, y_2} \max_{\xi_2 \sim (\bar{\mu}_2, \Gamma_2)} \mathbb{E}[V_3]$$

subject to

$$e^T x_2 + y_2 = W_1$$

$$V_1 = \min_{x_1, y_1} \max_{\xi_1 \sim (\bar{\mu}_1, \Gamma_1)} \mathbb{E}[V_2]$$

subject to

$$e^T x_1 + y_1 = W_0.$$

There is a very clear and simple economic interpretation of the dynamic portfolio rules of Theorem 4.2.1 and Corollary 4.2.1. The dynamic rule gives a myopic dynamic portfolio policy in the following sense. The single period optimal portfolio policy consists in setting a target excess wealth beyond the level which could be obtained by putting all the present wealth in the riskless asset: $r - W_0 R$. Dividing this excess target wealth by the optimal Sharpe ratio H^2 plus one (or a more complicated factor involving H and ϵ in the case of Theorem 4.2.1), one obtains the coefficient in the optimal rule. A similar formula is given in the previous results for the multi-period case. Note that each term $\frac{r - W_{t-1} R^{T-t+1}}{R^{T-t}}$ has the following economic meaning: the investor looks at the end of the current period t and sets the excess wealth target equal to

$$\frac{r}{R^{T-t}} - W_{t-1} R$$

which is exactly the discounted target wealth value at time $t+1$ minus the wealth that would be obtained if the current wealth W_t was kept in the riskless account for one period. If this number is equal to zero or is negative, then the final target can simply be achieved by investing the current wealth into the riskless asset for the rest of the horizon; hence the optimal position in risky assets would be zero for the remaining periods. If this excess target remains positive for all periods t, a constant multiple of it divided by the optimal period t Sharpe ratio H_t^2 (or another more complicated looking factor involving H and ϵ) plus one we have the optimal rule for each period. In other words, it is as if the investor is solving at each time period the following problem

$$\min_{x_t, y_t} \max_{\xi_t \sim (\bar{\mu}_t, \Gamma_t), \bar{\mu}_t \in U_{\bar{\mu}}^t} \mathbb{E}[\frac{r}{R^{T-t}} - \xi_t^T x_t - R y_t]_+$$

subject to

$$e^T x_t + y_t = W_{t-1}.$$

4.3 Conclusion

We derived closed form solutions for portfolio selection problems in single period and multiple period investment environments in this chapter. The results followed similar lines to those in the previous chapter, while the risk measure used here was the expected shortfall from a target, i.e. LPM_1, instead of the squared semi-deviation from a target. The results in this chapter reflect a more moderate risk aversion and we observe bolder investment decisions regarding risky assets, with each risky asset receiving twice the wealth allocation as compared to the investment decisions of the previous chapter. Portfolio rules in this chapter are derived under both distribution ambiguity of asset returns and ambiguity of the mean return vector. The incorporation of a riskless asset both simplifies the portfolio rules, and reflects the preference of the investor to refrain from investing into risky assets. In multiple periods, the optimal portfolio rule is a myopic decision following the rule derived for the single period investment: If the target has not been reached, it is as if the investor is solving in each period the single period problem with some adjustments to the target. In the following chapter, we use $Value-at-Risk$ and $Conditional-Value-at-Riks$ as risk measures instead of lower partial moments. Lower partial moments incorporate a parameter expressing the expected return target of the investor. Since both $Value-at-Risk$ and $Conditional-Value-at-Riks$ lack such aspect on the gain side, we build models analogous to those in this chapter by adding minimum expected return constraints to formulations.

Chapter 5

Robust Portfolio Choice with CVaR and VaR under Distribution and Mean Return Ambiguity

The organization of this chapter is as follows: in Section 5.1, in the case of n risky assets and a riskless asset, under distribution ambiguity where the mean return vector μ and variance-covariance matrix Γ of risky assets are known, robust portfolios are computed, minimizing the worst-case CVaR risk measure under distribution uncertainty. In Section 5.2 the Mean Return Ambiguity issue is addressed, deriving closed-form optimal portfolio rules for worst-case CVaR (and worst-case VaR) under both distribution and ellipsoidal mean return ambiguity, in the existence of riskless and risky assets. In the case the riskless asset is not included in the portfolio problem, in Section 5.3 we derive in closed form the optimal portfolio choice robust against distribution and ellipsoidal mean return ambiguity without using a minimum mean return constraint, generalizing a result in [23] that considers the case of distribution ambiguity only, i.e., full confidence in the mean return estimate. This study is published in TOP - Journal of the Spanish Society of Statistics and Operations Research [69].

5.1 Minimizing Robust CVaR and VaR in the Presence of a Riskless Asset under Distribution Ambiguity

We work in a financial market with a riskless asset with return rate R in addition to n risky assets. We investigate robust solutions minimizing Conditional Value-at-Risk (CVaR) and Value-at-Risk (VaR) measures. The unit initial wealth is allocated into the riskless and risky assets so as to allow short positions, thus we can define the loss function as a function of the vector $x \in \mathbb{R}^n$ of allocations to n risky assets and the random vector ξ of return rates for these assets:

$$f(x,\xi) \;=\; -\left(x^T\xi + R\left(1 - x^T e\right)\right), \tag{5.1}$$

where e denotes an n-vector of all ones. For the calculation of CVaR and VaR, we use the method in [55], minimizing over γ the auxiliary function:

$$F_\theta(x,\gamma) \;=\; \gamma + \frac{1}{1-\theta}\mathbb{E}\left[\left(f(x,\xi) - \gamma\right)_+\right], \tag{5.2}$$

i.e., the θ-conditional value-at-risk is calculated as follows:

$$CVaR_\theta(x) \;=\; \min_{\gamma\in\mathbb{R}} F_\theta(x,\gamma), \tag{5.3}$$

where θ is the threshold probability level or the quantile parameter, which is generally taken to be in the interval $[0.95, 1)$. The convex set consisting of γ values that minimize F_θ contains the θ-value-at-risk, $VaR_\theta(x)$, which is the minimum value in the set.

The worst-case CVaR, when ξ may assume a distribution from the set $\mathbb{D} = \{\pi|\mathbb{E}_\pi[\xi] = \mu, \mathrm{Cov}_\pi[\xi] = \Gamma \succ 0\}$ (i.e., the set of all distributions with (known and/or trusted) mean μ and covariance Γ), is defined as follows:

$$\begin{aligned} RCVaR_\theta(x) &= \max_{\pi\in\mathbb{D}} CVaR_\theta(x) && (5.4)\\ &= \max_{\pi\in\mathbb{D}} \min_{\gamma\in\mathbb{R}} F_\theta(x,\gamma). && (5.5)\end{aligned}$$

For convenience, in the rest of the chapter we replace occurrences of the sup operator with max since in all cases we consider the sup is attained.

We assume μ is not a multiple of e, as usual. Let us define the excess mean return $\tilde{\mu} = \mu - Re$, and the highest attainable Sharpe ratio in the market $H = \tilde{\mu}^T \Gamma^{-1} \tilde{\mu}$; see [28]. The following theorem gives an explicit solution of the portfolio choice problem of minimizing worst-case CVaR under a minimum mean return constraint

$$\min_{x \in \mathbb{R}^n : (\mu - Re)^T x \geq d - R} RCVaR_\theta \left(x \right), \tag{5.6}$$

where $d \in \mathbb{R}_+$ is a minimum target mean return parameter. We assume the minimum mean target return is larger than the riskless return, i.e., $d > R$.

Theorem 5.1.1. *For $\theta \geq \frac{H}{H+1}$, the problem (5.6) admits the optimal portfolio rule*

$$x^* = \frac{d - R}{H} \Gamma^{-1} \tilde{\mu}.$$

For $\theta < \frac{H}{H+1}$, the problem (5.6) is unbounded.

Proof. Since the set of distributions $\mathbb{D}$ is convex and the function $\gamma + \frac{1}{1-\theta} \mathbb{E}\left[\left(f\left(x, \xi\right) - \gamma \right)_+ \right]$ is convex in γ for every ξ, we can interchange supremum and minimum; see Theorem 2.4 of [70]. More precisely, one considers first the problem $\min_{\gamma \in \mathbb{R}} \max_{\pi \in \mathbb{D}} F_\theta \left(x, \gamma \right)$ and one finds a unique optimal solution as we shall do below. Then, using convexity of $\mathbb{D}$ and the convexity of the function $\gamma + \frac{1}{1-\theta} \mathbb{E}\left[\left(f\left(x, \xi\right) - \gamma \right)_+ \right]$ in γ one invokes Theorem 2.4 of [70] that allows interchange of min and max. We carry out this sequence of operations below:

$$\begin{aligned}
\min_{\gamma \in \mathbb{R}} \max_{\pi \in \mathbb{D}} F_\theta \left(x, \gamma \right) &= \min_{\gamma \in \mathbb{R}} \gamma + \frac{1}{1-\theta} \max_{\pi \in \mathbb{D}} \mathbb{E}\left[\left(-x^T \xi - R + R x^T e - \gamma \right)_+ \right] && (5.7) \\
&= \min_{\gamma \in \mathbb{R}} \gamma + \frac{1}{1-\theta} \max_{\xi \sim (\mu, \Gamma)} \mathbb{E}\left[\left(-R - \gamma - x^T \left(\xi - Re \right) \right)_+ \right] && (5.8) \\
&= \min_{\gamma \in \mathbb{R}} \gamma + \frac{1}{1-\theta} \max_{\eta \sim (\nu, \sigma^2)} \mathbb{E}\left[\left(-R - \gamma - \eta \right)_+ \right] && (5.9) \\
&= \min_{\gamma \in \mathbb{R}} \gamma + \frac{1}{1-\theta} \frac{-R - \gamma - \nu + \sqrt{\sigma^2 + \left(-R - \gamma - \nu \right)^2}}{2} && (5.10)
\end{aligned}$$

Maximum operators in (5.8) and (5.9) are equivalent, since by a slight modification of Lemma 2.4 in [23] we can state the equivalence of following sets of univariate random distributions:

$$\begin{aligned}\mathbb{D}_1 &= \left\{x^T(\xi - Re) : \xi \sim (\mu, \Gamma)\right\}\\ \mathbb{D}_2 &= \left\{\eta : \eta \sim \left(\nu, \sigma^2\right)\right\},\end{aligned}$$

where $\nu = x^T(\mu - Re)$, and $\sigma^2 = x^T\Gamma x$. Equality (5.10) follows by Lemma 2.2 of [23]. Then $RCVaR_\theta(x)$ is the minimization of the following function over γ:

$$h_x(\gamma) = \gamma + \frac{1}{1-\theta}\frac{-R-\gamma-\nu+\sqrt{\sigma^2+(R+\gamma+\nu)^2}}{2}.$$

γ_x^* minimizing $h_x(\gamma)$ can be found by equating the first derivative to zero:

$$h'_x(\gamma) = 1 + \frac{1}{1-\theta}\frac{-1+(R+\gamma+\nu)\left(\sigma^2+(R+\gamma+\nu)^2\right)^{-\frac{1}{2}}}{2}.$$

Hence, we have

$$2\theta - 1 = (R+\gamma+\nu)/\sqrt{\sigma^2+(R+\gamma+\nu)^2},$$

(recall that $\theta > 0.5$ so that $2\theta - 1 > 0$) or, equivalently,

$$(2\theta-1)^2\left(\sigma^2+(R+\gamma+\nu)^2\right) = (R+\gamma+\nu)^2,$$

or,

$$(2\theta-1)^2\sigma^2 = (R+\gamma+\nu)^2\left(1-(2\theta-1)^2\right)$$

whence we obtain

$$\begin{aligned}R+\gamma+\nu &= \frac{(2\theta-1)\sigma}{\sqrt{4\theta-4\theta^2}}\\ \gamma_x^* &= -R-\nu+\frac{(2\theta-1)\sigma}{2\sqrt{\theta}\sqrt{1-\theta}},\end{aligned}$$

since $h_x(\gamma)$ is a convex function as is verified immediately using the second derivative test:

$$h''_x(\gamma) = \frac{1}{2(1-\theta)}\left(\sigma^2+(R+\gamma+\nu)^2\right)^{-\frac{3}{2}}\sigma^2 \geq 0.$$

The minimum value of $h_x(\gamma)$ being known, we can calculate the min max value:

$$\begin{aligned}
\min_{\gamma\in\mathbb{R}}\max_{\pi\in\mathbb{D}} F_\theta(x,\gamma) &= \min_{\gamma\in\mathbb{R}} h_x(\gamma) \\
&= h_x(\gamma_x^*) \\
&= -R-\nu+\frac{(2\theta-1)\,\sigma}{2\sqrt{\theta}\sqrt{1-\theta}}+\frac{1}{1-\theta}\frac{\frac{-(2\theta-1)\sigma}{2\sqrt{\theta}\sqrt{1-\theta}}+\sqrt{\sigma^2+\frac{(2\theta-1)^2}{4\theta(1-\theta)}\sigma^2}}{2} \\
&= -R-\nu+\frac{(2\theta-1)\,\sigma}{2\sqrt{\theta}\sqrt{1-\theta}}+\frac{1}{1-\theta}\frac{\sigma}{2}\left(\frac{-2\theta+1+1}{2\sqrt{\theta}\sqrt{1-\theta}}\right) \\
&= -R-\nu+\frac{(2\theta-1)\,\sigma}{2\sqrt{\theta}\sqrt{1-\theta}}+\frac{1}{1-\theta}\frac{\sigma\,(1-\theta)}{2\sqrt{\theta}\sqrt{1-\theta}} \\
&= -R-\nu+\frac{2\theta\sigma}{2\sqrt{\theta}\sqrt{1-\theta}} \\
&= -R-\nu+\frac{\sqrt{\theta}}{\sqrt{1-\theta}}\sigma \\
&= -R-(\mu-Re)^T x+\frac{\sqrt{\theta}}{\sqrt{1-\theta}}\sqrt{x^T\Gamma x}. \qquad (5.11)
\end{aligned}$$

Hence, by Theorem 2.4 of [70] we have

$$RCVaR_\theta(x) = -R-(\mu-Re)^T x+\frac{\sqrt{\theta}}{\sqrt{1-\theta}}\sqrt{x^T\Gamma x}.$$

Minimizing the above expression for worst-case CVaR under the minimum mean return constraint, the robust optimal portfolio selection can be found[1]. Using a non-negative multiplier λ the Lagrange function is

$$L(x,\lambda) = -R-(\mu-Re)^T x+\frac{\sqrt{\theta}}{\sqrt{1-\theta}}\sqrt{x^T\Gamma x}+\lambda\left(d-R-(\mu-Re)^T x\right).$$

The first-order conditions are:

$$-\mu+Re+\frac{\tau\Gamma x}{\sqrt{x^T\Gamma x}}-\lambda(\mu-Re)=0,$$

[1]It is a simple exercise to show that in the absence of the minimum mean return constraint the portfolio position in the ith risky asset tends to $\pm\infty$ depending on the sign of the ith component of $\tilde{\Gamma}^{-1}\tilde{\mu}$.

where we defined $\tau \equiv \frac{\sqrt{\theta}}{\sqrt{1-\theta}}$ for convenience. We make the supposition that $x \neq 0$ and define $\sigma \equiv \sqrt{x^T \Gamma x}$. We get the candidate solution

$$x_c = \frac{\sigma(\lambda+1)}{\tau}\Gamma^{-1}\tilde{\mu}.$$

Now, using the identity $x_c^T \Gamma x_c = \sigma^2$, we obtain the equation

$$(\lambda+1)^2 = \frac{\tau^2}{H},$$

which implies that $\lambda = \frac{\tau}{\sqrt{H}} - 1$. Under the condition $\frac{\tau}{\sqrt{H}} \geq 1$ we ensure $\lambda \geq 0$. Now, utilizing the constraint which we assumed would be tight, the resulting equation yields (after substituting for x_c and λ)

$$\sigma = \frac{d-R}{\sqrt{H}},$$

which is positive under the assumption $d > R$ (H is positive by positive definiteness of Γ and the assumption that μ is not a multiple of e). Now, substitute the above expressions for σ and λ into x_c, and the desired expression is obtained after evident simplification. If $\frac{\tau}{\sqrt{H}} < 1$ then the problem (5.6) is unbounded by the convex duality theorem since it is always feasible. □

While our proof is similar to the proof in Chen *et al.* [23] in essence, their proof is faulty because their argument for exchanging max and min relies on a result of Zhu and Fukushima [57] which is valid only for discrete distributions. Our setting here, like that of Chen *et al.* [23] is not confined to discrete distributions. Hence, a different justification is needed for exchanging max and min. The Value-At-Risk measure can also be calculated using the auxiliary function (5.2):

$$VaR_\theta(x) = \arg\min_{\gamma \in \mathbb{R}} F_\theta(x,\gamma).$$

The worst-case value-at-risk is now calculated as follows:

$$\begin{aligned} RVaR_\theta(x) &= \max_{\pi \in \mathbb{D}} VaR_\theta(x) \\ &= \max_{\pi \in \mathbb{D}} \arg\min_{\gamma \in \mathbb{R}} F_\theta(x,\gamma) \\ &= \arg\min_{\gamma \in \mathbb{R}} \max_{\pi \in \mathbb{D}} F_\theta(x,\gamma) \\ &= \arg\min_{\gamma \in \mathbb{R}} h_x(\gamma) \\ &= \gamma_x^*. \end{aligned}$$

Again, a change of the order of operators makes this calculation possible by Theorem 2.4 of [70]. With a similar approach to that followed for CVaR, we pose the problem of minimizing the robust VaR:

$$\min_{x\in\mathbb{R}^n} -R - x^T\mu + Rx^Te + \frac{(2\theta-1)}{2\sqrt{\theta}\sqrt{1-\theta}}\sqrt{x^T\Gamma x}$$

subject to

$$(\mu - Re)^T x \geq d - R.$$

We obtain a replica of the previous result in this case. Hence the proof, which is identical, is omitted.

Theorem 5.1.2. *For $\theta \geq \frac{1}{2} + \frac{1}{2}\frac{\sqrt{H}}{\sqrt{H+1}}$, the problem of minimizing robust VaR under distribution ambiguity and a minimum mean return restriction admits the optimal portfolio rule*

$$x^* = \frac{d-R}{H}\Gamma^{-1}\tilde{\mu}.$$

For $\theta < \frac{1}{2} + \frac{1}{2}\frac{\sqrt{H}}{\sqrt{H+1}}$, the problem is unbounded.

In both results stated above, the optimal portfolios are mean-variance efficient. We plot the critical thresholds for θ given in Theorem 5.1.1 and Theorem 5.1.2 in Figure 5.1. Both thresholds tend to one as H goes to infinity and the threshold curve for robust VaR dominates that of robust CVaR. Hence, robust VaR leads to aggressive portfolio behavior in a larger interval for θ than robust CVaR. In other words, one has to choose a larger confidence level θ for robust VaR compared to robust CVaR in order to make an optimal portfolio choice. Thus we can affirm that robust VaR is more conservative than robust CVaR.

Based on the above results, it is straightforward to derive the equations of the robust CVaR efficient frontier and the robust VaR efficient frontier. Both robust frontiers are the straight lines governed by the equation

$$d = \frac{\tau}{\tau - \sqrt{H}}R + \frac{\sqrt{H}}{\tau - \sqrt{H}}f, \tag{5.12}$$

where for $f = RCVaR$ and $\tau = \frac{\sqrt{\theta}}{\sqrt{1-\theta}}$ we obtain the efficient frontier for robust CVaR; and for $f = RVaR$ and $\tau = \frac{(2\theta-1)}{2\sqrt{\theta}\sqrt{1-\theta}}$ we get the robust VaR efficient frontier.

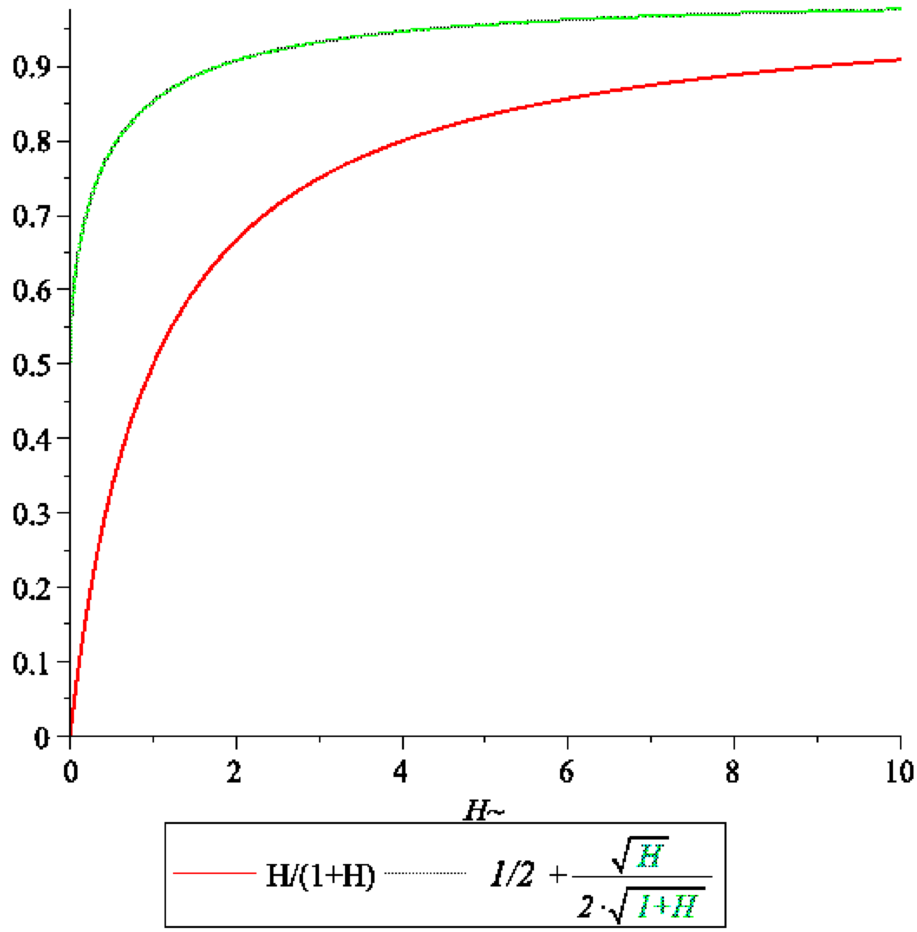

Figure 5.1: The critical thresholds for Robust CVaR and Robust VaR. The upper curve is the threshold value curve for robust VaR.

5.2 Robust CVaR and VaR in the Presence of Riskless Asset under Distribution and Mean Return Ambiguity

We consider now the problem of choosing a portfolio $x \in \mathbb{R}^n$ that minimizes the function

$$RRCVaR_\theta(x) = \max_{\pi \in \mathbb{D}, \bar{\mu} \in U_{\bar{\mu}}} CVaR_\theta(x) \tag{5.13}$$

or equivalently

$$RRCVaR_\theta(x) = \max_{\bar{\mu} \in U_{\bar{\mu}}} \max_{\pi \in \mathbb{D}} \min_{\gamma \in \mathbb{R}} F_\theta(x, \gamma) \tag{5.14}$$

where we define the ellipsoidal uncertainty set $U_{\bar{\mu}} = \{\bar{\mu} | \|\Gamma^{-1/2}(\bar{\mu} - \mu^{\mathrm{nom}})\|_2 \leq \sqrt{\epsilon}\}$ for the mean return denoted $\bar{\mu}$, where μ^{nom} denotes a nominal mean return vector which can be considered as the available estimate of mean return. The parameter ϵ acts as a measure of confidence in the mean return estimate. We consider now the problem of choosing a portfolio $x \in \mathbb{R}^n$ that minimizes the function

$$RRCVaR_\theta(x) = \max_{\bar{\mu} \in U_{\bar{\mu}}} \max_{\pi \in \mathbb{D}} \min_{\gamma \in \mathbb{R}} F_\theta(x, \gamma) \tag{5.15}$$

subject to

$$(\bar{\mu} - Re)^T x \geq d - R.$$

As in the previous section we define $\mathbb{D} = \{\pi | \mathbb{E}_\pi[\xi] = \bar{\mu}, \mathrm{Cov}_\pi[\xi] = \Gamma \succ 0\}$. Considering the inner max min problem over $\pi \in \mathbb{D}$ and $\gamma \in \mathbb{R}$, respectively, as in the proof of Theorem 5.1.1 of previous section while we keep $\bar{\mu}$ fixed, we arrive at the inner problem for the objective function (recall equation (5.11)) that we can transform immediately:

$$\begin{aligned} \max_{\bar{\mu} \in U_{\bar{\mu}}} RCVaR_\theta(x) &= \max_{\bar{\mu} \in U_{\bar{\mu}}} -R - (\bar{\mu} - Re)^T x + \frac{\sqrt{\theta}}{\sqrt{1-\theta}} \sqrt{x^T \Gamma x} \\ &= -R - (\mu^*)^T x + \sqrt{\epsilon}\sqrt{x^T \Gamma x} + \frac{\sqrt{\theta}}{\sqrt{1-\theta}} \sqrt{x^T \Gamma x}, \end{aligned}$$

where the last equality follows using a well-known transformation result in robust optimization (see e.g., [6]), and $\mu^* = \mu^{\mathrm{nom}} - Re$.

Using the same transformation on the minimum mean return constraint as well, we obtain the second-order conic problem

$$\min_{x \in \mathbb{R}^n} -R - x^T \mu^* + (\sqrt{\epsilon} + \frac{\sqrt{\theta}}{\sqrt{1-\theta}})\sqrt{x^T \Gamma x}$$

subject to

$$(\mu^*)^T x - \sqrt{\epsilon}\sqrt{x^T \Gamma x} \geq d - R,$$

where $d > R$.

Theorem 5.2.1. *Under the Slater constraint qualification,*

1. *If* $\theta > \frac{(\sqrt{H}-\sqrt{\epsilon})^2}{(\sqrt{H}-\sqrt{\epsilon})^2+1}$, *and* $\epsilon < H$ *then the problem (5.15) admits the optimal portfolio rule*
$$x^* = \frac{d-R}{(\sqrt{H}-\sqrt{\epsilon})\sqrt{H}} \Gamma^{-1} \tilde{\mu}.$$
2. *If* $\epsilon = H$, *the problem is unbounded.*
3. *If* $\epsilon > H$, *the problem is infeasible.*

Proof. Under the Slater constraint qualification the Karush-Kuhn-Tucker optimality conditions are both necessary and sufficient. Using a non-negative multiplier λ we have the Lagrange function

$$L(x, \lambda) = -R - x^T \mu^* + (\sqrt{\epsilon} + \frac{\sqrt{\theta}}{\sqrt{1-\theta}})\sqrt{x^T \Gamma x} + \lambda \left(d - R - (\mu^*)^T x + \sqrt{\epsilon}\sqrt{x^T \Gamma x} \right).$$

Going through the usual steps as in the proof of Theorem 5.1.1 under the supposition that σ (defined as $\sqrt{x^T \Sigma x}$) is a finite non-zero positive number, we have the candidate solution:

$$x_c = \frac{\sigma(\lambda+1)}{\tau + \sqrt{\epsilon}(\lambda+1)} \Gamma^{-1} \mu^*.$$

From the identity $x_c^T \Gamma x_c = \sigma^2$ we obtain the quadratic equation in λ:

$$(H-\epsilon)\lambda^2 + (2H - 2\epsilon - 2\sqrt{\epsilon}\tau)\lambda + H - \epsilon - 2\sqrt{\epsilon}\tau - \tau^2 = 0$$

with the two roots

$$-\frac{\tau + \sqrt{H} + \sqrt{\epsilon}}{\sqrt{H} + \sqrt{\epsilon}}, -\frac{-\tau + \sqrt{H} - \sqrt{\epsilon}}{\sqrt{H} - \sqrt{\epsilon}}.$$

The left root cannot be positive, so it is discarded. The right root is positive if $H > \epsilon$ and $\tau \geq \sqrt{H} - \sqrt{\epsilon}$. Now, returning to the conic constraint which is assumed to be tight we obtain an equation in σ after substituting for λ in the right root, and solving for σ using straightforward algebraic simplification we get

$$\sigma = \frac{d - R}{\sqrt{H} - \sqrt{\epsilon}},$$

which is positive provided $H > \epsilon$. Now, the expression for x^* is obtained after substitution for σ and λ into x_c and evident simplification.

Part 2 is immediate from the result of Part 1.

For Part 3, if $\epsilon \geq H$, our hypothesis of a finite positive non-zero σ is false, in which case the only possible value for σ is zero, achieved at the zero risky portfolio which is infeasible. □

The problem of minimizing the robust VaR under distribution and mean return ambiguity in the presence of a minimum target mean return constraint is posed as

$$\min_{x \in \mathbb{R}^n} -R - x^T \mu^* + \left(\sqrt{\epsilon} + \frac{(2\theta - 1)}{2\sqrt{\theta}\sqrt{1-\theta}}\right)\sqrt{x^T \Gamma x}$$

subject to

$$(\mu^*)^T x - \sqrt{\epsilon}\sqrt{x^T \Gamma x} \geq d - R.$$

Again, we obtain a result similar to the previous theorem in this case. The proof is a verbatim repetition of the proof of the previous theorem, hence omitted.

Theorem 5.2.2. *Under the Slater constraint qualification,*

1. *If* $\theta > \frac{1}{2} + \frac{\sqrt{H}-\sqrt{\epsilon}}{2\sqrt{1+(\sqrt{H}-\sqrt{\epsilon})^2}}$, *and* $\epsilon < H$ *then the problem of minimizing the robust VaR under distribution and mean return ambiguity in the presence of a minimum target mean return constraint admits the optimal portfolio rule*

$$x^* = \frac{d - R}{(\sqrt{H} - \sqrt{\epsilon})\sqrt{H}}\Gamma^{-1}\tilde{\mu}.$$

2. *If* $\epsilon = H$, *the problem is unbounded.*

3. If $\epsilon > H$, the problem is infeasible.

Notice that we obtain identical and mean-variance efficient portfolio rules for both CVaR and VaR under distribution and mean return ambiguity. Furthermore, the optimal portfolio rules reduce to those of Theorem 5.1.1 and Theorem 5.1.2, respectively, for $\epsilon = 0$, the case of distribution ambiguity only.

The robust CVaR and VaR efficient frontiers are straight lines given by the equation:

$$d = \frac{\tau}{\tau - \sqrt{H} + \sqrt{\epsilon}} R + \frac{\sqrt{H} - \sqrt{\epsilon}}{\tau - \sqrt{H} + \sqrt{\epsilon}} f,$$

where for $f = RRCVaR$ and $\tau = \frac{\sqrt{\theta}}{\sqrt{1-\theta}}$ we obtain the efficient frontier for robust CVaR; and for $f = RRVaR$ and $\tau = \frac{(2\theta-1)}{2\sqrt{\theta}\sqrt{1-\theta}}$ we get the robust VaR efficient frontier.

Comparing the above results and efficient frontier to the results of the previous section and to the efficient frontier equation (5.12) for the case of ambiguity distribution only, we notice that the effect of introducing mean return ambiguity in addition to distribution ambiguity has the effect of replacing $\sqrt{H}$ by $\sqrt{H} - \sqrt{\epsilon}$. More precisely, the mean return ambiguity decreases the optimal Sharpe ratio of the market viewed by the investor. The investor can form an optimal portfolio in the risky assets as long as his/her confidence in the mean return vector is not too low, i.e., his ϵ does not exceed the optimal Sharpe ratio of the market.

The efficient frontier line for the case of robust portfolios in the face of both distribution and mean return ambiguity is less steep than the efficient robust portfolios for distribution ambiguity only. This simple fact can be verified by direct computation:

$$\frac{\sqrt{H} - \sqrt{\epsilon}}{\tau - \sqrt{H} + \sqrt{\epsilon}} - \frac{\sqrt{H}}{\tau - \sqrt{H}} = -\frac{\tau\sqrt{\epsilon}}{(\tau - \sqrt{H})(\tau - \sqrt{H} + \sqrt{\epsilon})} < 0.$$

The d-intercept for the former is also smaller than for the latter as can be easily seen. We provide a numerical illustration in Figure 5.2 with $H = 0.47222$, $\epsilon = 0.4$, $R = 1.01$ and $\theta = 0.95$. The efficient portfolios robust to distribution ambiguity are represented by the steeper line. It is clear that the incremental effect of mean return ambiguity and robustness is to render the investor more risk averse and more cautious.

5.3 Robust CVaR and VaR without the Riskless Asset under Distribution and Mean Return Ambiguity

We consider now the problems of the previous section as treated in [23], i.e., without the riskless asset and without the minimum mean return constraint since we do not need this restriction to obtain an optimal portfolio rule. In that case we are dealing with the loss function:

$$f(x,\xi) \quad = \quad -x^T\xi \tag{5.16}$$

and the auxiliary function

$$F_\theta(x,\gamma) \quad = \quad \gamma + \frac{1}{1-\theta}\mathbb{E}\left[\left(-x^T\xi-\gamma\right)_+\right]. \tag{5.17}$$

In [23] Theorem 2.9 the authors solve the problem

$$\min_{x\in\mathbb{R}^n:e^Tx=1}\ \max_{\pi\in\mathbb{D}} CVaR_\theta(x)$$

in closed form. We shall now attack the problem under the assumption that the mean returns are subject to errors that we confine to the ellipsoid: $U_{\bar{\mu}} = \{\bar{\mu} | \|\Gamma^{-1/2}(\bar{\mu}-\mu^{\text{nom}})\|_2 \leq \sqrt{\epsilon}\}$ for the mean return denoted $\bar{\mu}$, where μ^{nom} denotes a nominal mean return vector as in the previous section. That is, we are interested in solving

$$\min_{x\in\mathbb{R}^n:e^Tx=1}\ \max_{\bar{\mu}\in U_{\bar{\mu}}}\max_{\pi\in\mathbb{D}} CVaR_\theta(x).$$

Using partially the proof of Theorem 2.9 in [23], and Theorem 2.4 of [70] as in the previous section, we have for arbitrary $\bar{\mu}$

$$\max_{\pi\in\mathbb{D}} CVaR_\theta(x) = \sqrt{\frac{\theta}{1-\theta}}\sqrt{x^T\Gamma x} - x^T\bar{\mu},$$

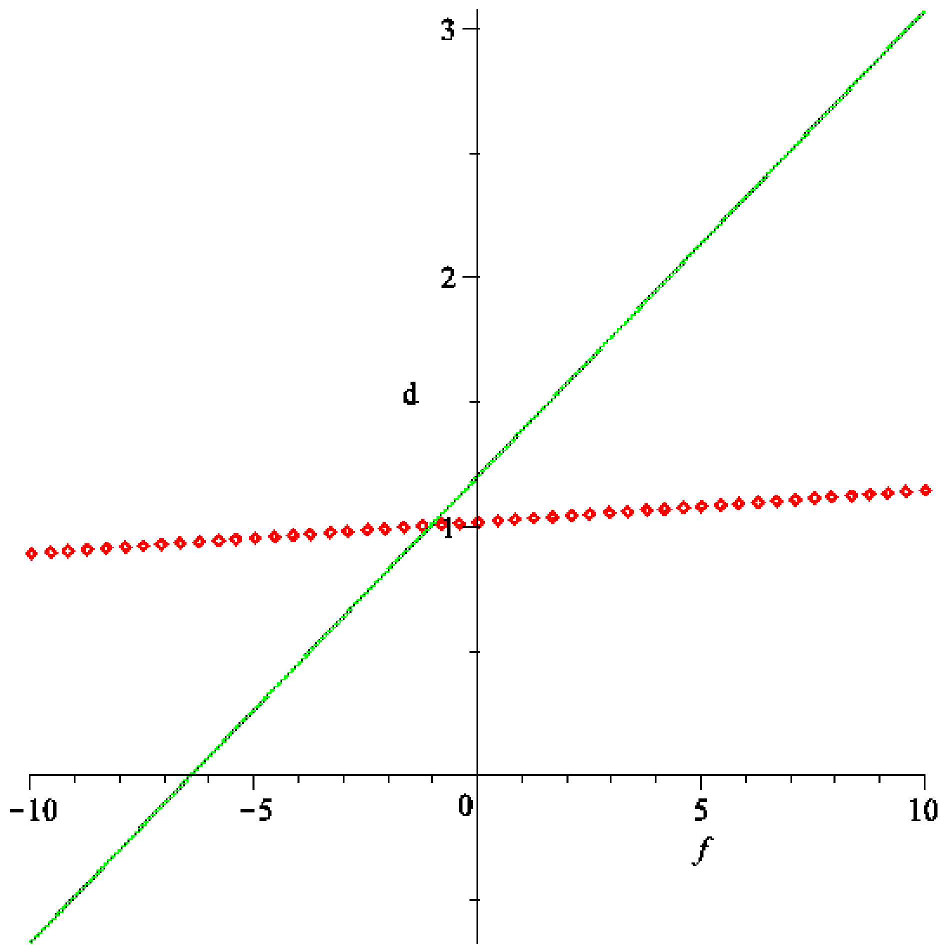

Figure 5.2: The efficient frontier lines for Robust CVaR and Robust VaR for $H = 0.47222$, $\epsilon = 0.4$, $R = 1.01$ and $\theta = 0.95$. The steeper line corresponds to distribution ambiguity case while the point line corresponds to distribution and mean return ambiguity case.

and the maximum is attained at

$$\frac{2\theta-1}{2\sqrt{\theta(1-\theta)}}\sqrt{x^T\Gamma x}-x^T\bar{\mu}$$

which happens to be equal to the robust VaR under distribution ambiguity. Therefore, we obtain

$$\max_{\bar{\mu}\in U_{\bar{\mu}}}\max_{\pi\in\mathbb{D}} CVaR_\theta(x)=\sqrt{\frac{\theta}{1-\theta}}\sqrt{x^T\Gamma x}-x^T\mu^{\text{nom}}+\sqrt{\epsilon}\sqrt{x^T\Gamma x}.$$

Now, we are ready to process the problem

$$\min_{x\in\mathbb{R}^n:e^Tx=1}\sqrt{\frac{\theta}{1-\theta}}\sqrt{x^T\Gamma x}-x^T\mu^{\text{nom}}+\sqrt{\epsilon}\sqrt{x^T\Gamma x}.$$

From the first-order conditions we obtain

$$x=\frac{\sigma}{\tau+\sqrt{\epsilon}}\Gamma^{-1}(\mu^{\text{nom}}+\lambda e)$$

where $\sigma\equiv(x^T\Gamma x)^{1/2}$, $\tau=\sqrt{\theta/(1-\theta)}$ and λ is the Lagrange multiplier. Using the equation $\sigma^2=x^T\Gamma x$ we obtain the quadratic equation:

$$C\lambda^2+2B\lambda+A-(\tau+\sqrt{\epsilon})^2=0$$

where $A=(\mu^{\text{nom}})^T\Gamma^{-1}\mu^{\text{nom}}$, $C=e^T\Gamma^{-1}e$, $B=e^T\Gamma^{-1}\mu^{\text{nom}}$. We solve for λ under the condition

$$C(\tau+\sqrt{\epsilon})^2>AC-B^2$$

(note that $AC-B^2>0$ by Cauchy-Schwarz inequality):

$$\lambda=\frac{-B+\sqrt{B^2-AC+C(\tau+\sqrt{\epsilon})^2}}{C}.$$

We discard the root $\frac{-B-\sqrt{B^2-AC+C(\tau+\sqrt{\epsilon})^2}}{C}$ because it leads to a negative value for σ; see the expression for σ below. In this case the dual problem is infeasible and the primal is unbounded since it cannot be infeasible. The condition $C(\tau+\sqrt{\epsilon})^2>AC-B^2$ is equivalent to

$$\epsilon>(\sqrt{A-B^2/C}-\sqrt{\theta/(1-\theta)})^2,$$

(we do not allow $\epsilon = (\sqrt{A - B^2/C} - \sqrt{\theta/(1-\theta)})^2$ since it results in a σ that grows without bound, hence the problem is unbounded). Using the above expression for λ in the equation $e^T x = 1$ we solve for σ to get

$$\sigma = \frac{\tau + \sqrt{\epsilon}}{\sqrt{B^2 - AC + C(\tau + \sqrt{\epsilon})^2}}.$$

Substituting λ and σ to the expression for x obtained from the first-order conditions we obtain the solution

$$x^* = \frac{1}{\sqrt{C(\tau+\sqrt{\epsilon})^2 - \Delta}} \Gamma^{-1} \mu^{\text{nom}} + \left[-\frac{B}{C\sqrt{C(\tau+\sqrt{\epsilon})^2 - \Delta}} + \frac{1}{C} \right] \Gamma^{-1} e,$$

where we have defined $\Delta = AC - B^2$ for simplicity. The portfolio problem for robust VaR under distribution and ellipsoidal mean return ambiguity is resolved similarly. In fact, the only change is in the definition of τ. Therefore, we have proved the following result.

Theorem 5.3.1.

1. *The distribution and mean return ambiguity robust CVaR portfolio choice, i.e., the solution to problem*

$$\min_{x \in \mathbb{R}^n : e^T x = 1} \max_{\bar{\mu} \in U_{\bar{\mu}}} \max_{\pi \in \mathbb{D}} CVaR_\theta(x)$$

   is given by

$$x^* = \frac{1}{\sqrt{C(\tau+\sqrt{\epsilon})^2 - \Delta}} \Gamma^{-1} \mu^{nom} + \left[-\frac{B}{C\sqrt{C(\tau+\sqrt{\epsilon})^2 - \Delta}} + \frac{1}{C} \right] \Gamma^{-1} e, \tag{5.18}$$

   where $\Delta = AC - B^2$, $A = (\mu^{nom})^T \Gamma^{-1} \mu^{nom}$, $C = e^T \Gamma^{-1} e$, $B = e^T \Gamma^{-1} \mu^{nom}$, *and* $\tau = \sqrt{\theta/(1-\theta)}$ *provided that* $\epsilon > (\sqrt{A - B^2/C} - \tau)^2$. *If* $\epsilon \leq (\sqrt{A - B^2/C} - \tau)^2$ *the problem is unbounded.*

2. *The distribution and mean return ambiguity robust VaR portfolio choice, i.e., the solution to problem*

$$\min_{x \in \mathbb{R}^n : e^T x = 1} \max_{\bar{\mu} \in U_{\bar{\mu}}} \frac{2\theta - 1}{2\sqrt{\theta(1-\theta)}} \sqrt{x^T \Gamma x} - x^T \bar{\mu}$$

is given by

$$x^* = \frac{1}{\sqrt{C(\tau+\sqrt{\epsilon})^2 - \Delta}}\Gamma^{-1}\mu^{nom} + \left[-\frac{B}{C\sqrt{C(\tau+\sqrt{\epsilon})^2-\Delta}} + \frac{1}{C}\right]\Gamma^{-1}e. \tag{5.19}$$

where $\Delta = AC - B^2$, $A = (\mu^{nom})^T\Gamma^{-1}\mu^{nom}$, $C = e^T\Gamma^{-1}e$, $B = e^T\Gamma^{-1}\mu^{nom}$, *and* $\tau = \frac{2\theta-1}{2\sqrt{\theta(1-\theta)}}$ *provided that* $\epsilon > (\sqrt{A - B^2/C} - \tau)^2$. *If* $\epsilon \leq (\sqrt{A - B^2/C} - \tau)^2$ *the problem is unbounded.*

Notice that the optimal portfolios are mean-variance efficient. Using the data in [23] we plot the optimal CVaR function $\sqrt{\frac{\theta}{1-\theta}}\sqrt{x^T\Gamma x} - x^T\mu^{\text{nom}} + \sqrt{\epsilon}\sqrt{x^T\Gamma x}$ evaluated at x^* as a function of ϵ in Figure 5.2 for two different values of θ. As the confidence parameter ϵ increases (i.e., confidence in the mean return estimate diminishes), the optimal robust CVaR increases, which implies an increase in risk. A similar behavior occurs for robust VaR optimal value function. Interestingly the optimal robust CVaR and VaR increase for constant ϵ when the quantile parameter θ increases.

When the uncertainty set $U_{\bar{\mu}}$ is reduced to a single point $\mu^{\text{nom}} = \mu$, i.e., for $\epsilon = 0$ we obtain precisely the result in Theorem 2.9 of [23]:

Corollary 5.3.1.

1. *The distribution ambiguity robust CVaR portfolio choice, i.e., the solution to problem*

$$\min_{x\in\mathbb{R}^n : e^Tx=1}\ \max_{\pi\in\mathbb{D}} CVaR_\theta(x)$$

is given by

$$x^* = \frac{1}{\sqrt{C\tau^2 - \Delta}}\Gamma^{-1}\mu + \left[-\frac{B}{C\sqrt{C\tau^2-\Delta}} + \frac{1}{C}\right]\Gamma^{-1}e. \tag{5.20}$$

where $\Delta = AC - B^2$, $A = \mu^T\Gamma^{-1}\mu$, $C = e^T\Gamma^{-1}e$, $B = e^T\Gamma^{-1}\mu$, *and* $\tau = \sqrt{\theta/(1-\theta)}$ *provided that* $\theta/(1-\theta) > A - B^2/C$. *If* $\theta/(1-\theta) \leq A - B^2/C$ *the problem is unbounded.*

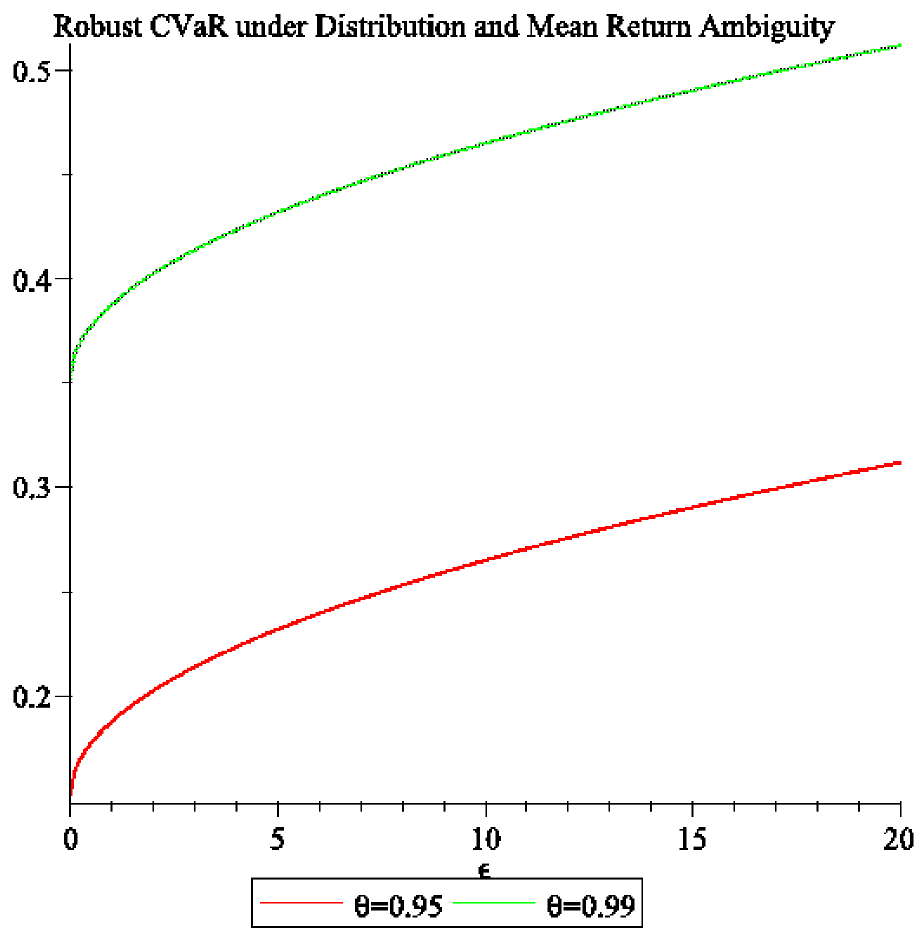

Figure 5.3: The behavior of robust CVaR as a function of ϵ in the case without the riskless asset. The upper curve is for $\theta = 0.99$ and the lower curve is for $\theta = 0.95$.

2. *The distribution ambiguity robust VaR portfolio choice, i.e., the solution to problem*

$$\min_{x \in \mathbb{R}^n : e^T x = 1} \frac{2\theta - 1}{2\sqrt{\theta(1-\theta)}} \sqrt{x^T \Gamma x} - x^T \mu$$

is given by

$$x^* = \frac{1}{\sqrt{C\tau^2 - \Delta}} \Gamma^{-1}\mu + \left[-\frac{B}{C\sqrt{C\tau^2 - \Delta}} + \frac{1}{C} \right] \Gamma^{-1} e. \tag{5.21}$$

where $\Delta = AC - B^2$, $A = \mu^T \Gamma^{-1} \mu$, $C = e^T \Gamma^{-1} e$, $B = e^T \Gamma^{-1} \mu$, *and* $\tau = \frac{2\theta-1}{2\sqrt{\theta(1-\theta)}}$ *provided that* $(\frac{2\theta-1}{2\sqrt{\theta(1-\theta)}})^2 > A - B^2/C$. *If* $(\frac{2\theta-1}{2\sqrt{\theta(1-\theta)}})^2 \leq A - B^2/C$ *the problem is unbounded.*

5.4 Conclusion

In this chapter, we studied the choice on worst-case Value-at-Risk and worst-case Conditional-Value-at-Risk minimizing portfolio. Since the risk measures discussed in this chapter do not reflect the investor's expectation on mean return, we added a constraint limiting from below the expected return of the portfolio selection. In the existence of n risky assets and a riskless asset, under both distribution ambiguity and distribution ambiguity combined with ambiguity in the mean return vector, we derived closed-form portfolio rules minimizing the two risk measures. In this chapter, the problems are restricted to static settings. We concluded the chapter with the extension of a previous result, in a setting without a riskless asset and the minimum expected return constraint. We derived the optimal portfolio choice robust against distribution and ellipsoidal mean return uncertainty. Our closed-form portfolio choice in this setting extends and generalizes the previous results derived under distribution uncertainty with fixed expected return vector.

Chapter 6

Investor Avoidance from Risk as Uniform Portfolio Becomes Optimal

In a market with N assets, the uncertainty involving asset return measure Q on $\mathbb{R}^N$ is known to be inside an open ball around a known nominal measure P. The definition of the ball of radius κ, $\mathcal{B}_\kappa(P)$, is based on the Kantorovich/Wasserstein distance of degree p:

$$d_p(P_1, P_2) \quad = \quad \inf \left\{ \left(\int_{\mathbb{R}^N \times \mathbb{R}^N} \|x-y\|_p^p d\pi(x,y) \right)^{\frac{1}{p}} : proj_1(\pi) = P_1, proj_2(\pi) = P_2 \right\},$$

as defined for two measures P_1 and P_2 in $\mathbb{R}^N$. π is a measure on $\mathbb{R}^{\mathbb{N}} \times \mathbb{R}^{\mathbb{N}}$, with marginals P_1 and P_2. The infimum is known to be attained by such a measure, and the minimizing π is called the optimal transportation plan. Based on this uncertainty set, a robust portfolio selection problem arises, and is defined by Pflug et al. [5] as follows:

$$\inf_{w \in \mathbb{R}^N} \sup_{Q \in \mathcal{B}_\kappa(P)} \mathcal{R}(\langle X^Q, w \rangle) \tag{6.1}$$

$$s.t. \quad \langle w, \mathbb{1} \rangle = 1. \tag{6.2}$$

The risk functional $\mathcal{R} : L^p(\Omega, \sigma, \mu) \to \mathbb{R}$ is convex, version independent and is required to have dual structure:

$$\mathcal{R}(X) = \max \{\mathbb{E}(XZ) - R(Z) : Z \in L^q\}.$$

It is proved that as the uncertainty radius κ increases, the optimal solution of the above problem either equals or otherwise converges to the uniform portfolio $w^u = \frac{1}{N}\mathbb{1}$. In this problem, the investor is forced to invest all his wealth to the market of N assets. We know his decision given that he invests all his wealth into the risky market, but this is not really informative on whether he is really interested in entering the market when uncertainty and risk increases. By introducing a riskless asset, we try the assess the inclination of the investor to avoid the risky assets as uncertainty increases to the level at which optimal (risky) investment converges to the uniform portfolio of risky assets. The problem is redefined as follows:

$$\inf_{\substack{w \in \mathbb{R}^N \\ s \in \mathbb{R}}} \sup_{Q \in \mathcal{B}_\kappa(P)} \mathcal{R}(\langle X^Q, w \rangle + \eta s) \tag{6.3}$$

$$s.t. \;\; \langle w, \mathbb{1} \rangle + s = 1. \tag{6.4}$$

We first fix the amount invested into the riskless asset, s, to prove that the optimal choice for the remaining wealth and assets converges to the uniform portfolio $w^{u,s} = \frac{1-s}{N}\mathbb{1}$). Fixing s at a given value, we define an inner problem as follows:

$$\inf_{w \in \mathbb{R}^N} \sup_{Q \in \mathcal{B}_\kappa(P)} \mathcal{R}(\langle X^Q, w \rangle + \eta s) \tag{6.5}$$

$$s.t. \quad \langle w, \mathbb{1} \rangle = 1 - s. \tag{6.6}$$

6.1 Convergence of Risky Investment to Uniform Portfolio

To prove (approximate) optimality of $w^{u,s}$ for (6.5)-(6.6), some preliminary results are necessary. Proofs follow the lines of Pflug et al. [5], nevertheless restatement is necessary due to the change in objective function, hence problem structure.

Lemma 6.1.1. *Let $\mathcal{R} : L^p(\Omega, \sigma, \mu) \to \mathbb{R}$ be a convex, version independent risk measure with dual representation, as defined above, $1 \leq p < \infty$. Let $\frac{1}{p} + \frac{1}{q} = 1$. Then,*

$$\begin{aligned} &\left|\mathcal{R}\left(\left\langle X^{P_1}, w\right\rangle + \eta\left(1 - e^T w\right)\right) - \mathcal{R}\left(\left\langle X^{P_2}, w\right\rangle + \eta\left(1 - e^T w\right)\right)\right| \\ &\qquad \leq \sup_{Z: R(Z) < \infty} \|Z\|_{L^q} \|w\|_q \, d_p\left(P_1, P_2\right) \end{aligned} \tag{6.7}$$

for arbitrary measures P_1, P_2 on $\mathbb{R}^N$.

Proof. Let π be the minimizer of the Kantorivich distance between P_1 and P_2, and let $Y : (\Omega, \sigma, \mu) \to \mathbb{R}^N \times \mathbb{R}^N$ with the image measure on $\mathbb{R}^N \times \mathbb{R}^N$ equal to π. Let the projections of Y on the first and second components be X^{P_1} and X^{P_2}, respectively. Note that the image measure of X^{P_i} is P_i, $i = 1, 2$. When Z is chosen as the maximizer in the definition of $\mathcal{R}$ for the first random variable:

$$\mathbb{E}\left[\left(\left\langle X^{P_1}, w\right\rangle + \eta\left(1 - e^T w\right)\right) Z\right] - R(Z), \tag{6.8}$$

we have that:

$$
\begin{aligned}
& \mathcal{R}\left(\left\langle X^{P_1}, w\right\rangle + \eta\left(1 - e^T w\right)\right) - \mathcal{R}\left(\left\langle X^{P_2}, w\right\rangle + \eta\left(1 - e^T w\right)\right) && (6.9)\\
\leq\ & \mathbb{E}\left[\left(\left\langle X^{P_1}, w\right\rangle + \eta\left(1 - e^T w\right)\right) Z\right] - R(Z) - \mathbb{E}\left[\left(\left\langle X^{P_2}, w\right\rangle + \eta\left(1 - e^T w\right)\right) Z\right] + R(Z) && \\
=\ & \mathbb{E}\left[\left\langle X^{P_1} - X^{P_2}, w\right\rangle Z\right] && (6.10)\\
\leq\ & \mathbb{E}\left[\left|\left\langle X^{P_1} - X^{P_2}, w\right\rangle Z\right|\right] && (6.11)\\
\leq\ & \left(\mathbb{E}\left[\left|\left\langle X^{P_1} - X^{P_2}, w\right\rangle\right|^p\right]\right)^{1/p} \left(\mathbb{E}\left[|Z|^q\right]\right)^{1/q} && (6.12)\\
=\ & \left(\int_\Omega \left|\left\langle X^{P_1} - X^{P_2}, w\right\rangle\right|^p d\mu\right)^{1/p} \left(\int_\Omega |Z|^q \, d\mu\right)^{1/q} && (6.13)\\
=\ & \|Z\|_{L^q} \left(\int_\Omega \left|\left\langle X^{P_1} - X^{P_2}, w\right\rangle\right|^p d\mu\right)^{1/p} && (6.14)\\
=\ & \|Z\|_{L^q} \left(\int_\Omega \left|\sum_{n=1}^N \left(X_n^{P_1} - X_n^{P_2}\right) w_n\right|^p d\mu\right)^{1/p} && (6.15)\\
\leq\ & \|Z\|_{L^q} \left(\int_\Omega \left(\sum_{n=1}^N \left|\left(X_n^{P_1} - X_n^{P_2}\right) w_n\right|\right)^p d\mu\right)^{1/p} && (6.16)\\
\leq\ & \|Z\|_{L^q} \left(\int_\Omega \left(\left(\sum_{n=1}^N \left|X_n^{P_1} - X_n^{P_2}\right|^p\right)^{1/p} \left(\sum_{n=1}^N |w_n|^q\right)^{1/q}\right)^p d\mu\right)^{1/p} && (6.17)\\
=\ & \|Z\|_{L^q} \left(\int_\Omega \left(\sum_{n=1}^N \left|X_n^{P_1} - X_n^{P_2}\right|^p\right) \|w\|_q^p d\mu\right)^{1/p} && (6.18)\\
=\ & \|Z\|_{L^q} \|w\|_q \left(\int_\Omega \sum_{n=1}^N \left|X_n^{P_1} - X_n^{P_2}\right|^p d\mu\right)^{1/p} && (6.19)\\
=\ & \|Z\|_{L^q} \|w\|_q \left(\int_{\mathbb{R}\times\mathbb{R}} \sum_{n=1}^N \left|x_n^1 - x_n^2\right|^p d\pi\left(x^1, x^2\right)\right)^{1/p} && (6.20)\\
=\ & \|Z\|_{L^q} \|w\|_q \, d_p\left(P_1, P_2\right) && (6.21)\\
\leq\ & \sup_{Z: R(Z) < \infty} \|Z\|_{L^q} \|w\|_q \, d_p\left(P_1, P_2\right) && (6.22)
\end{aligned}
$$

(6.12) follows from the application of Hölder's Inequality on the random variable function spaces, and (6.17) follows from Hölder's Inequality on $\mathbb{R}^N$. The arguments are based on a Z specific to minimizing $\mathcal{R}\left(\left\langle X^{P_1}, w\right\rangle + \eta\left(1 - e^T w\right)\right)$, but in the final step, taking supremum over the set of possible Z, this dependence is alleviated. Thus, repeating the arguments for $\mathcal{R}\left(\left\langle X^{P_2}, w\right\rangle + \eta\left(1 - e^T w\right)\right) - \mathcal{R}\left(\left\langle X^{P_1}, w\right\rangle + \eta\left(1 - e^T w\right)\right)$, we reach the result. □

We will now show that for every measure P and corresponding random variable X^P, there exist Q / X^Q at which the bound proved in Lemma (6.1.1) is attained, i.e. the inequality holds with inequality. We begin with the case $p \in (1, \infty)$.

Proposition 6.1.1. *Let the risk functional $\mathcal{R}$ be as defined above. Let $1 < p < \infty$ and q such that $\frac{1}{p} + \frac{1}{q} = 1$. Let P be a probability measure on $\mathbb{R}^N$, and assume that:*

$$\|Z\|_{L^q} = C \text{ for all } Z \in \bigcup_{X \in L^p} \partial\mathcal{R}(X) \text{ with } R(Z) < \infty. \tag{6.23}$$

Then it holds that for every $\kappa > 0$ and every $w \in \mathbb{R}^N$, there is a measure Q on $\mathbb{R}^N$ such that $d_p(P, Q) = \kappa$ and

$$\left|\mathcal{R}\left(\left\langle X^Q, w\right\rangle + \eta\left(1 - e^T w\right)\right) - \mathcal{R}\left(\left\langle X^P, w\right\rangle + \eta\left(1 - e^T w\right)\right)\right| = C\kappa\|w\|_q, \tag{6.24}$$

i.e., the bound of Lemma 6.1.1 holds with equality.

Proof. Fix a $Z \in \partial\mathcal{R}\left(\left\langle X^P, w\right\rangle + \eta\left(1 - e^T w\right)\right)$ with $R(Z) < \infty$. Let X^Q be a random variable defined as follows:

$$X_n^Q = X_n^P + c_1(n)\,|w_n|^{\frac{q}{p}} \text{ with} \tag{6.25}$$

$$c_1(n) = \frac{sign(w_n)\, sign(Z) c_2}{\|w\|_q^q}\,|Z|^{\frac{q}{p}} \tag{6.26}$$

for all $n : 1 \leq n \leq N$ and $c_2 > 0$. Setting $c_1 = |c_1(n)|$, it follows that:

$$c_1^p\,|w_n|^q = \left|X_n^Q - X_n^P\right|^p, \; \forall n : 1 \leq n \leq N. \tag{6.27}$$

Also,

$$\left|\sum_{n=1}^N w_n\left(X_n^Q - X_n^P\right)\right|^p = \left|\sum_{n=1}^N w_n c_1(n)\,|w|^{\frac{q}{p}}\right|^p = c_1^p\left|\sum_{n=1}^N |w_n|^q\right|^p = c_1^p\|w\|_q^{pq} = c_2^p\,|Z|^q. \tag{6.28}$$

Note that the choice of the parameter $c_2 > 0$ determines the distance $d_p(P, Q)$ of the image measure Q of X^Q to P, i.e., bigger values yield a bigger distance, and for every $\kappa > 0$, there is a $c_2 > 0$ such that $d_p(P, Q) = \kappa$ for the respective image measure Q. Again, selecting Z maximizing

$$\mathbb{E}\left[\left(\left\langle X^Q, w\right\rangle + \eta\left(1 - e^T w\right)\right) Z\right] - R(Z), \tag{6.29}$$

the result can be obtained as follows:

$$\left|\mathcal{R}\left(\left\langle X^Q, w\right\rangle + \eta\left(1 - e^T w\right)\right) - \mathcal{R}\left(\left\langle X^P, w\right\rangle + \eta\left(1 - e^T w\right)\right)\right| \tag{6.30}$$

$$\geq \mathbb{E}\left(\left\langle X^Q, w\right\rangle + \eta\left(1 - e^T w\right)\right) - R(Z) - \mathbb{E}\left(\left\langle X^Q, w\right\rangle + \eta\left(1 - e^T w\right)\right) + R(Z) \tag{6.31}$$

$$= \mathbb{E}\left(\left\langle X^Q - X^P, w\right\rangle Z\right) \tag{6.32}$$

$$= \mathbb{E}\left(\left(\sum_{n=1}^{N}\left(X_n^Q - X_n^P\right) w_n\right) Z\right) \tag{6.33}$$

$$= \mathbb{E}\left(\left|\sum_{n=1}^{N}\left(X_n^Q - X_n^P\right) w_n\right| |Z|\right) \tag{6.34}$$

$$= \left(\mathbb{E}\left|\sum_{n=1}^{N}\left(X_n^Q - X_n^P\right) w_n\right|^p\right)^{1/p} \left(\mathbb{E}\,|Z|^q\right)^{1/q} \tag{6.35}$$

$$= \left(\mathbb{E}\left(\sum_{n=1}^{N}\left|X_n^Q - X_n^P\right| |w_n|\right)^p\right)^{1/p} \|Z\|_{L^q} \tag{6.36}$$

$$= \left(\mathbb{E}\left(\sum_{n=1}^{N}\left|X_n^Q - X_n^P\right|^p\right)\right)^{1/p} \|w\|_q \|Z\|_{L^q} \tag{6.37}$$

$$= \|Z\|_{L^q}\|w\|_q \left(\int_\Omega \sum_{n=1}^{N}\left|X_n^Q - X_n^P\right|^p d\mu\right)^{1/p} \tag{6.38}$$

$$= \|Z\|_{L^q}\|w\|_q \kappa \tag{6.39}$$

□

Transitions to (6.34) and (6.36) are possible due to the incorporation of $sign(w_n)$ and $sign(Z)$ in $X^Q - X^P$, which makes all terms in the sum non-negative. (6.35) and (6.37) are applications of Hölder's Inequality, where conditions for equality in Hölder's are assured by the selection $X^Q - X^P$. (6.35) is possible since $\left|\sum_{n=1}^{N}\left(X_n^Q - X_n^P\right) w_n\right|^p$ is equal to $|Z|^q$, when constant multipliers set aside. Similarly, (6.27) implies the condition for equality on transition to (6.37).

A similar result on the tightness of the bound given in Lemma 6.1.1 follows for the case $p = 1$.

Proposition 6.1.2. *Let $\mathcal{R} : \mathcal{L}^1(\Omega, \sigma, \mu) \to \mathbb{R}$ be a convex, version independent risk measure as defined above. Assume:*

$$\|Z\|_{\mathcal{L}^\infty} = C, \ |Z| = C, \ or \ |Z| = 0 \tag{6.40}$$

almost everywhere, for all possible subgradients of $\mathbb{R}$. Then it holds that for every probability measure P on $\mathbb{R}^N$ and $\kappa > 0$, there is a measure Q on $\mathbb{R}^N$ such that $d_1(P, Q) = \kappa$ and

$$\left|\mathcal{R}\left(\left\langle X^P, w\right\rangle + \eta\left(1 - e^T w\right)\right) - \mathcal{R}\left(\left\langle X^Q, w\right\rangle + \eta\left(1 - e^T w\right)\right)\right| = C\|w\|_\infty \kappa. \tag{6.41}$$

Proof. Let

$$X_n^Q = X_n^P + c_1(n) \tag{6.42}$$

$$c_1(n) = \begin{cases} sign(w_n)\, sign(Z)\, c_2, & |w_n| = \|w\|_\infty \\ 0, & \text{otherwise} \end{cases} \tag{6.43}$$

where we define $sign(0) = 0$.
Take Z that is a maximizer in $\mathcal{R}$ at $\left\langle X^P, w\right\rangle + \eta\left(1 - e^T w\right)$. Then:

$$\mathcal{R}\left(\left\langle X^Q, w\right\rangle + \eta\left(1 - e^T w\right)\right) - \mathcal{R}\left(\left\langle X^P, w\right\rangle + \eta\left(1 - e^T w\right)\right) \tag{6.44}$$

$$\geq \mathbb{E}\left(\left\langle X^Q - X^P, w\right\rangle Z\right) \tag{6.45}$$

$$= \mathbb{E}\left(\left|\left\langle X^Q - X^P, w\right\rangle Z\right|\right) \tag{6.46}$$

$$= \mathbb{E}\left(\left|\left\langle X^Q - X^P, w\right\rangle\right|\right) \|Z\|_{\mathcal{L}^\infty} \tag{6.47}$$

$$= \mathbb{E}\left(\sum_{n=1}^{N} \left|X_n^Q - X_n^P\right|\right) \|w\|_\infty \|Z\|_{\mathcal{L}^\infty} \tag{6.48}$$

$$= d_k(P, Q)\, \|w\|_\infty \|Z\|_{\mathcal{L}^\infty} \tag{6.49}$$

$$= \kappa \|w\|_\infty \|Z\|_{\mathcal{L}^\infty}. \tag{6.50}$$

□

We have posed the problem as follows:

$$\begin{array}{cl} \inf\limits_{\substack{w \in \mathbb{R}^N \\ s \in \mathbb{R}}} \sup\limits_{Q \in \mathcal{B}_\kappa(P)} & \mathcal{R}(\langle X^Q, w\rangle + \eta s) \\ \text{s.t.} & \langle w, \mathbb{1}\rangle + s = 1. \end{array}$$

Propositions 6.1.1 and 6.1.2, and the choice of distribution Q used in the proofs, indicate that:

$$\mathcal{R}(\langle X^Q, w\rangle + \eta s) - \mathcal{R}(\langle X^P, w\rangle + \eta s) = C\kappa \|w\|_q,$$

combined with Lemma 6.1.1, we obtain the following result:

$$\begin{aligned} \sup_{\bar{Q} \in \mathcal{B}_\kappa(P)} \mathcal{R}(\langle X^{\bar{Q}}, w\rangle + \eta s) &= \mathcal{R}(\langle X^Q, w\rangle + \eta s) \\ &= \mathcal{R}(\langle X^P, w\rangle + \eta s) + C\kappa \|w\|_q. \end{aligned}$$

With this result the worst case/supremum operator in the problem definition is eliminated:

$$\begin{array}{cl} \inf\limits_{\substack{w \in \mathbb{R}^N \\ s \in \mathbb{R}}} & \mathcal{R}(\langle X^P, w\rangle + \eta s) + C\kappa \|w\|_q \\ \text{s.t.} & \langle w, \mathbb{1}\rangle + s = 1, \end{array}$$

where fixing s, the problem can be solved in two steps:

$$\begin{array}{l} \inf\limits_{s \in \mathbb{R}} \inf\limits_{w \in \mathbb{R}^N} \mathcal{R}(\langle X^P, w\rangle + \eta s) + C\kappa \|w\|_q \\ \quad s.t.\ \langle w, \mathbb{1}\rangle = 1 - s. \end{array}$$

We will first investigate the optimal solutions of the inner problem:

$$\inf_{w \in \mathbb{R}^N} \mathcal{R}(\langle X^P, w\rangle + \eta s) + C\kappa \|w\|_q \tag{6.51}$$

$$s.t.\ \langle w, \mathbb{1}\rangle = 1 - s. \tag{6.52}$$

The two lemmas that follow will be used in the derivation of optimality results:

Lemma 6.1.2. *$\mathcal{R}(\langle X^P, w^1\rangle + \eta s) - \mathcal{R}(\langle X^P, w^2\rangle + \eta s) \leq C\|w^1 - w^2\|_q \mathbb{E}(\|X^P\|_p^p \mathbb{1}_{\{Z\neq 0\}})^{\frac{1}{p}}$ for all $w^1, w^2 \in \mathbb{R}^N$, and $Z \in \partial\mathcal{R}(\langle X^P, w^1\rangle + \eta s)$.*

Proof. Taking arbitrary $Z \in \partial\mathcal{R}(\langle X^P, w^1\rangle + \eta s)$,

$$\begin{aligned}
\mathcal{R}(\langle X^P, w^1\rangle + \eta s) &- \mathcal{R}(\langle X^P, w^2\rangle + \eta s) \\
&\leq \mathbb{E}[(\langle X^P, w^1\rangle + \eta s)Z] - R(Z) - \mathbb{E}[(\langle X^P, w^1\rangle + \eta s)Z] + R(Z) \\
&= \mathbb{E}[\langle X^P, w^1 - w^2\rangle Z] \\
&= \mathbb{E}[\langle X^P, w^1 - w^2\rangle \mathbb{1}_{\{Z\neq 0\}} Z] \\
&\leq \mathbb{E}[|\langle X^P, w^1 - w^2\rangle \mathbb{1}_{\{Z\neq 0\}} Z|] \\
&\leq (\mathbb{E}[|\langle X^P, w^1 - w^2\rangle \mathbb{1}_{\{Z\neq 0\}}|^p])^{\frac{1}{p}} \|Z\|_{L^q} \qquad (6.53) \\
&= (\mathbb{E}[|\langle X^P, w^1 - w^2\rangle|^p \mathbb{1}_{\{Z\neq 0\}}])^{\frac{1}{p}} \|Z\|_{L^q} \\
&\leq (\mathbb{E}[|\langle X^P, w^1 - w^2\rangle|^p \mathbb{1}_{\{Z\neq 0\}}])^{\frac{1}{p}} \|Z\|_{L^q} \\
&\leq (\mathbb{E}[(\|X^P\|_p \|w^1 - w^2\|_q)^p \mathbb{1}_{\{Z\neq 0\}}])^{\frac{1}{p}} \|Z\|_{L^q} \qquad (6.54) \\
&= (\mathbb{E}[\|X^P\|_p^p \mathbb{1}_{\{Z\neq 0\}}])^{\frac{1}{p}} \|w^1 - w^2\|_q \|Z\|_{L^q} \\
&= (\mathbb{E}[\|X^P\|_p^p \mathbb{1}_{\{Z\neq 0\}}])^{\frac{1}{p}} \|w^1 - w^2\|_q C. \qquad (6.55)
\end{aligned}$$

Inequality in (6.53) follows due to Hölder's Inequality applied on functions $(\langle X^P, w^1 - w^2\rangle \mathbb{1}_{\{Z\neq 0\}} \in L^p)$ and $Z \in L^q$. Inequality in (6.54) follows due to Hölder's Inequality applied to vectors in $\mathbb{R}^N$. The result, (6.55), follows due to our assumption that $\|Z\|_{L^q} = C$ on $\partial\mathcal{R}(\langle X^P, w^1\rangle + \eta s)$. □

Lemma 6.1.3. *For fixed $s \in \mathbb{R}$, the uniform portfolio $w^{u,s}$ is no worse than solutions in a set B, i.e.:*

$$\mathcal{R}(\langle X^P, w^{u,s}\rangle + \eta s) + C\|w^{u,s}\|_q \kappa \leq \mathcal{R}(\langle X^P, w\rangle + \eta s) + C\|w\|_q \kappa,$$

for all $w \in B$, if

$$\kappa \geq \frac{\|w - w^{u,s}\|_q}{\|w\|_q - \|w^{u,s}\|_q} \mathbb{E}(\|X^P\|_p^p \mathbb{1}_{\{Z\neq 0\}})^{\frac{1}{p}},$$

for all $w \in B$.

Proof. $\kappa \geq \frac{\|w-w^{u,s}\|_q}{\|w\|_q-\|w^u_s\|_q}\mathbb{E}(\|X^P\|_p^p\mathbb{1}_{\{Z\neq 0\}})^{\frac{1}{p}}$ implies, by multiplying both sides by $C(\|w\|_q - \|w^{u,s}\|_q)$:

$$\begin{aligned} C\kappa(\|w\|_q - \|w^{u,s}\|_q) &\geq C(\|w\|_q - \|w^{u,s}\|_q)\frac{\|w-w^{u,s}\|_q}{\|w\|_q - \|w^{u,s}\|_q}\mathbb{E}(\|X^P\|_p^p\mathbb{1}_{\{Z\neq 0\}})^{\frac{1}{p}} \\ &= C\|w-w^{u,s}\|_q\mathbb{E}(\|X^P\|_p^p\mathbb{1}_{\{Z\neq 0\}})^{\frac{1}{p}} \\ &\geq \mathcal{R}(\langle X^P, w^{u,s}\rangle + \eta s) - \mathcal{R}(\langle X^P, w\rangle + \eta s), \end{aligned}$$

where the last inequality follows by Lemma 6.1.2. Regrouping terms above, we have:

$$\mathcal{R}(\langle X^P, w\rangle + \eta s) + C\kappa\|w\|_q \geq \mathcal{R}(\langle X^P, w^{u,s}\rangle + \eta s) + C\kappa\|w^{u,s}\|_q,$$

for all $w \in B$ ($\|w\|_q \neq \|w\|_q$ is assumed, which holds true for $w \neq w^{u,s}$ on a hyperplane of fixed $\langle w, \mathbb{1}\rangle = \langle w, \mathbb{1}\rangle$, since $w^{u,s}$ uniquely minimizes $\|\cdot\|_q$ on that hyperplane). □

We are now ready to present optimality results for the inner problem (6.5)-(6.6).

Proposition 6.1.3. *Given fixed $s \in \mathbb{R}$, and considering $N \geq 2$,*

1. For $p = 1$, $w^{u,s}$ is the optimal solution to (6.51)-(6.52) if $\kappa \geq \kappa^*$, where:

$$\kappa^* = \begin{cases} (N-1)\mathbb{E}(\|X^P\|_1\mathbb{1}_{\{Z\neq 0\}}) & \text{if } s \neq 1 \\ \mathbb{E}(\|X^P\|_1\mathbb{1}_{\{Z\neq 0\}} & \text{if } s = 1. \end{cases}$$

2. For $p = 2$, the optimal portfolio for (6.51)-(6.52) lies in $\{w \in \mathbb{R}^N : \|w - w^{u,s}\|_q < D\}$ if:

$$\kappa \geq \left(\left(\frac{(1-s)^2}{ND^2}+1\right)^{\frac{1}{2}} + \frac{|1-s|}{\sqrt{N}D}\right)\mathbb{E}(\|X^P\|_2^2\mathbb{1}_{\{Z\neq 0\}})^{\frac{1}{2}}. \quad (6.56)$$

3. For $p \notin \{1, 2\}$,for every $\epsilon > 0$, there is a κ_ϵ such that for $\kappa \geq \kappa_\epsilon$ the optimal solution w^* of (6.51)-(6.52) is inside $\{w \in \mathbb{R}^N : \|w - w^{u,s}\|_q < \epsilon\}$.

Proof. Case $s = 1$, $1 \le p < \infty$:

In this case, (6.52) turns into $\langle w, \mathbb{1} \rangle = 0$, and $w^{u,s} = 0$, hence $\|w - w^{u,s}\|_q = \|w\|_q - \|w^{u,s}\|_q = \|w\|_q$. By Lemma 6.1.3,

$$\kappa \;\; \ge \;\; \mathbb{E}(\|X^P\|_p^p \mathbb{1}_{\{Z \neq 0\}})^{\frac{1}{p}}$$

implies

$$\kappa \;\; \ge \;\; \frac{\|w - w^{u,s}\|_q}{\|w\|_q - \|w^{u,s}\|_q} \mathbb{E}(\|X^P\|_p^p \mathbb{1}_{\{Z \neq 0\}})^{\frac{1}{p}}$$

for all $w \in \mathbb{R}^N$, $w \neq w^{u,s}$ hence optimality of $w^{u,s}$.

Case $s > 1$, $p = 1$:

In this case, (6.52) turns into $\langle w, \mathbb{1} \rangle = 1 - s$, where $1 - s < 0$, and $w^{u,s} = \frac{1-s}{N}$. For any w such that $\langle w, \mathbb{1} \rangle = 1 - s$, $w \neq w^{u,s}$, we let $n^* = \arg\max_{1 \le n \le N} \left| w_n - \frac{1-s}{N} \right|$. If $w_{n^*} < \frac{1-s}{N}$, $\|w - w^{u,s}\|_\infty = \left| w_{n^*} - \frac{1-s}{N} \right| = \|w\|_\infty - \|w^{u,s}\|_\infty$. Otherwise, $w_{n*} > \frac{1-s}{N}$, $w_{n^*} = \max_{1 \le n \le N} w_n$, and

$$\begin{aligned} \min_{1 \le n \le N} w_n \;\; &\le \;\; \frac{1-s}{N} - \frac{w_{n^*} - \frac{1-s}{N}}{N-1} \\ &= \;\; \frac{1-s}{N} - \frac{\|w - w^{u,s}\|_\infty}{N-1}. \end{aligned}$$

Since $\min_{1 \le n \le N} w_n \ge -\|w\|_\infty$, $\frac{1-s}{N} - \frac{\|w - w^{u,s}\|_\infty}{N-1} \ge -\|w\|_\infty$, and noting that $\frac{1-s}{N} = -\|w^{u,s}\|_\infty$, we have:

$$\|w\|_\infty - \|w^{u,s}\|_\infty \;\; \ge \;\; \frac{\|w - w^{u,s}\|_\infty}{N-1},$$

equivalently:

$$\frac{\|w - w^{u,s}\|_\infty}{\|w\|_\infty - \|w^{u,s}\|_\infty} \;\; \le \;\; N - 1.$$

Then if $\kappa \ge (N-1)\mathbb{E}(\|X^P\|_1 \mathbb{1}_{\{Z \neq 0\}})$, then $\kappa \ge \frac{\|w - w^{u,s}\|_\infty}{\|w\|_\infty - \|w^{u,s}\|_\infty} \mathbb{E}(\|X^P\|_1 \mathbb{1}_{\{Z \neq 0\}})$ for all $w \neq w^{u,s}$ such that $\langle w, \mathbb{1} \rangle = 1 - s$, and $w^{u,s}$ is optimal to (6.51)-(6.52).

Case $s < 1$, $p = 1$:

In this case, $\langle w, \mathbb{1}\rangle = 1 - s > 0$, and $w^{u,s} = \frac{1-s}{N}$. Similar to the case where $s > 1$, we set $n^* = \arg\max_{1\le n\le N} \left|w_n - \frac{1-s}{N}\right|$. If $w_{n^*} > \frac{1-s}{N}$, $\|w - w^{u,s}\|_\infty = \left(w_{n^*} - \frac{1-s}{N}\right) = \|w\|_\infty - \|w^{u,s}\|_\infty$. Otherwise, $w_{n^*} < \frac{1-s}{N}$, $w_{n^*} = \min_{1\le n\le N} w_n$, and

$$\begin{aligned} \max_{1\le n\le N} w_n &\ge \frac{1-s}{N} + \frac{w_{n^*} - \frac{1-s}{N}}{N-1} \\ &= \frac{1-s}{N} + \frac{\|w\|_\infty - \|w^{u,s}\|_\infty}{N-1} \end{aligned}$$

Since $\|w\|_\infty \ge \max_{1\le n\le N} w_n$, and $\|w^{u,s}\|_\infty = \frac{1-s}{N}$, we can write:

$$\begin{aligned} \|w\|_\infty - \|w^{u,s}\|_\infty &\ge \frac{\|w - w^{u,s}\|_\infty}{N-1} \\ N-1 &\ge \frac{\|w - w\|_\infty}{\|w\|_\infty - \|w^{u,s}\|_\infty}. \end{aligned}$$

Again, if $\kappa \ge (N-1)\mathbb{E}(\|X^P\|_1 \mathbb{1}_{\{Z\neq 0\}})$, the condition for Lemma 6.1.3 is satisfied in the feasible region and $w^{u,s}$ is optimal to (6.51)-(6.52).

Case $s \neq 1$, $p = 2$:

Let $f_2, \ldots, f_N$ unit vectors orthogonal to each other and $w^{u,s}$. Then a unique selection of $c_2, \ldots, c_N \in \mathbb{R}$ gives $w = w^{u,s} + \sum_{i=2}^N c_i f_i$ (since $w^{u,s} = \frac{1-s}{N}\mathbb{1}$, $f_2, \ldots, f_N$ are orthogonal to $w^{u,s}$, and $\langle w, \mathbb{1}\rangle = 1 - s$, $\langle w^{u,s}, \mathbb{1}\rangle = \langle w, \mathbb{1}\rangle$ implies that the coefficient of $w^{u,s}$ in w is equal to 1). Then:

$$\begin{aligned}
\frac{\|w-w^{u,s}\|_2}{\|w\|_2-\|w^{u,s}\|_2} &= \frac{\|w-w^{u,s}\|_2}{(\frac{(1-s)^2}{N}+\sum_{i=2}^N c_i^2)^{\frac{1}{2}}-\frac{|1-s|}{\sqrt{N}}} \\
&= \frac{\|w-w^{u,s}\|_2\left[(\frac{(1-s)^2}{N}+\sum_{i=2}^N c_i^2)^{\frac{1}{2}}+\frac{|1-s|}{\sqrt{N}}\right]}{\frac{(1-s)^2}{N}+\sum_{i=2}^N c_i^2-\frac{(1-s)^2}{N}} \\
&= \frac{\|w-w^{u,s}\|_2\left[(\frac{(1-s)^2}{N}+\sum_{i=2}^N c_i^2)^{\frac{1}{2}}+\frac{|1-s|}{\sqrt{N}}\right]}{\sum_{i=2}^N c_i^2} \\
&= \frac{\|w-w^{u,s}\|_2\left[(\frac{(1-s)^2}{N}+\sum_{i=2}^N c_i^2)^{\frac{1}{2}}+\frac{|1-s|}{\sqrt{N}}\right]}{\|w-w^{u,s}\|_2^2} \\
&= \frac{(\frac{(1-s)^2}{N}+\sum_{i=2}^N c_i^2)^{\frac{1}{2}}+\frac{|1-s|}{\sqrt{N}}}{\|w-w^{u,s}\|_2} \\
&= (\frac{(1-s)^2}{N\|w-w^{u,s}\|_2^2}+1)^{\frac{1}{2}}+\frac{|1-s|}{\sqrt{N}\|w-w^{u,s}\|_2}.
\end{aligned}$$

Defining the set B as $B := \{w \in \mathbb{R}^N : \|w-w^{u,s}\|_2 \geq D\}$, the above equality implies that inside the set B,

$$\frac{\|w-w^{u,s}\|_2}{\|w\|_2-\|w^{u,s}\|_2} \leq (\frac{(1-s)^2}{ND^2}+1)^{\frac{1}{2}}+\frac{|1-s|}{\sqrt{N}D}.$$

If the value of κ satisfies

$$\kappa \geq \left[(\frac{(1-s)^2}{ND^2}+1)^{\frac{1}{2}}+\frac{|1-s|}{\sqrt{N}D}\right]\mathbb{E}(\|X^P\|_2^2\mathbb{1}_{\{Z\neq 0\}})^{\frac{1}{2}}$$

then

$$\kappa \geq \frac{\|w-w^{u,s}\|_2}{\|w\|_2-\|w^{u,s}\|_2}\mathbb{E}(\|X^P\|_2^2\mathbb{1}_{\{Z\neq 0\}})^{\frac{1}{2}}$$

for $w \in B$, and by Lemma 6.1.3, a solution with better objective value than $w^{u,s}$ can only be inside $\{w \in \mathbb{R}^N : \|w-w^{u,s}\|_2 < D\}$.

Case $p \notin \{1,2\}$:

In this case, we show that for an increasing sequence κ_n, the optimal solution gets to fall inside a smaller neighborhood surrounding $w^{u,s}$ as $\kappa_n \longrightarrow \infty$. We define the set

$$A_n = \{w \in \mathbb{R}^N : \langle w, \mathbb{1}\rangle = 1 - s,$$
$$\mathcal{R}(\langle X^P, w^{u,s}\rangle + \eta s) + C\|w^{u,s}\|_q \kappa_n \geq \mathcal{R}(\langle X^P, w\rangle + \eta s) + C\|w\|_q \kappa_n\}.$$

Since C and κ_n are positive, $\mathcal{R}(\langle X^P, \cdot\rangle + \eta s)$ and $\|\cdot\|_q$ are convex functions, A_n is a convex set; and since $\|w\|_q - \|w^{u,s}\|_q \geq 0$ for $\langle w, \mathbb{1}\rangle = 1 - s$, $A_{n+1} \subseteq A_n$. $\cap_{n=1}^{\infty} A_n = \{w^{u,s}\}$,since for any $w \neq w^{u,s}$, $\langle w, \mathbb{1}\rangle = 1 - s$, and for sufficiently large κ_w,

$$\mathcal{R}(\langle X^P, w^{u,s}\rangle + \eta s) + C\|w^{u,s}\|_q \kappa_w \quad < \quad \mathcal{R}(\langle X^P, w\rangle + \eta s) + C\|w\|_q \kappa_w,$$

as $\|w\|_q > \|w^{u,s}\|_q$. A_n is bounded, since $\mathcal{R}(\langle X^P, w\rangle + \eta s) + C\|w\|_q \kappa_n \geq C\|w\|_q \kappa_n$ and both C and κ are positive.

One can easily show that A_n are closed sets, hence compact in this setting where boundedness is obvious.

Let $B_n^{\epsilon} = A_n \cap \{w \in \mathbb{R}^N : \|w - w^{u,s}\|_q \geq \epsilon\}$. $\cup_{n=1}^{\infty} B_n^{\epsilon} = \emptyset$, since $\cup_{n=1}^{\infty} A_n = \{w^{u,s}\}$ and $w^{u,s} \notin B_n^{\epsilon}$, $\forall n \in \mathbb{N}$. Compactness of B_n^{ϵ}, $n \in \mathbb{N}$, imply that there is $N^{\epsilon} \in \mathbb{N}$ such that $B_{N^{\epsilon}}^{\epsilon} = \emptyset$, that is, the optimal solution is inside $\{w \in \mathbb{R}^N : \|w - w^{u,s}\|_q < \epsilon\}$ for $\kappa \geq \kappa_{N^{\epsilon}}$. □

6.2 Division of Wealth between Riskless and Risky Assets

After the inner optimization problem (6.5)-(6.6) is solved, and (approximate) optimality of the uniform portfolio $w^{u,s}$ corresponding to the fixed value of s is proven, now it remains to solve for optimal value of s, that minimizes the risk measure $\mathcal{R}$, when combined with its corresponding risky and uniform portfolio $w^{u,s}$. Upto this point, the abstract definition of $\mathcal{R}$, limited to requirement of convexity, version independence and dual characterization properties, was sufficient for deriving optimal portfolio selection rules. However, for characterizing the optimality rules for the investment into the riskless asset, we will need to study specific risk measures. The Conditional Value-at-Risk (CVaR) measure is one possible choice, satisfying all assumptions on $\mathcal{R}$, and popular in risk minimizing optimization applications due to its desirable properties. Setting $p = 1$, with the risk measure specified as CVaR, the problem can be redefined as:

$$\begin{aligned} \inf_{\substack{w \in \mathbb{R}^N \\ s \in \mathbb{R}}} \sup_{Q \in \mathcal{B}_\kappa(P)} & CVaR_\alpha(-\langle X^Q, w\rangle - \eta s) \\ s.t. \quad & \langle w, \mathbb{1}\rangle + s = 1. \end{aligned}$$

If κ, hence the level of uncertainty in the risky assets, exceeds the threshold determined in Proposition 6.1.3, then the solution for the inner problem is known, and there remains to solve the outer problem:

$$\inf_{s \in \mathbb{R}} \sup_{Q \in \mathcal{B}_\kappa(P)} \quad CVaR_\alpha(-\langle X^Q, w^{u,s}\rangle - \eta s),$$

which based on our earlier results transforms to:

$$\inf_{s \in \mathbb{R}} \quad CVaR_\alpha(-\langle X^P, w^{u,s}\rangle - \eta s) + C\kappa \|w^{u,s}\|_\infty.$$

Plugging in $w^{u,s} = \frac{1-s}{N}\mathbb{1}$, we have:

$$\inf_{s \in \mathbb{R}} \quad CVaR_\alpha(-\tfrac{1-s}{N}\langle \mathbb{1}, X^P\rangle - \eta s) + C\kappa \tfrac{|1-s|}{N}.$$

If we let

$$f(s) = CVaR_\alpha(-\frac{1-s}{N}\langle \mathbb{1}, X^P\rangle - \eta s) + C\kappa \frac{|1-s|}{N},$$

then $f(1) = CVaR_\alpha(-\eta) = -\eta$. When $s < 1$, we can write:

$$\begin{aligned} f(s) &= CVaR_\alpha(-\frac{1-s}{N}\langle \mathbb{1}, X^P\rangle - \eta s) + C\kappa \frac{|1-s|}{N} \\ &= \frac{1-s}{N} CVaR_\alpha(-\langle \mathbb{1}, X^P\rangle) - \eta s + C\kappa \frac{|1-s|}{N}. \end{aligned}$$

When $s > 1$, we have:

$$\begin{aligned} f(s) &= CVaR_\alpha(-\frac{1-s}{N}\langle \mathbb{1}, X^P\rangle - \eta s) + C\kappa\frac{|1-s|}{N} \\ &= \frac{s-1}{N} CVaR_\alpha(\langle \mathbb{1}, X^P\rangle) - \eta s + C\kappa\frac{|1-s|}{N}. \end{aligned}$$

Denoting $CVaR_\alpha(-\langle \mathbb{1}, X^P\rangle)$ by γ^- and $CVaR_\alpha(\langle \mathbb{1}, X^P\rangle)$ by γ^+, we can summarize f as follows:

$$f(s) = \begin{cases} \frac{1-s}{N}\gamma^- - \eta s + C\kappa\frac{|1-s|}{N} & if\ s < 1 \\ \eta & if\ s = 1 \\ \frac{s-1}{N}\gamma^+ - \eta s + C\kappa\frac{|1-s|}{N} & if\ s > 1, \end{cases}$$

rearranging terms, we have:

$$f(s) = \begin{cases} s(-\frac{\gamma^-}{N} - \eta) + \frac{\gamma^-}{N} + C\kappa\frac{|1-s|}{N} & if\ s < 1 \\ \eta & if\ s = 1 \\ s(\frac{\gamma^+}{N} - \eta) - \frac{\gamma^+}{N} + C\kappa\frac{|1-s|}{N} & if\ s > 1. \end{cases}$$

Note that γ^+ and γ^- are constants given the nominal distribution P, the former represents average conditional loss for a uniform positive unit allocation to each risky asset, and the latter represents that for a negative allocation. Hence if this conditional value at risk (divided by N) dominates the sum of the riskless rate η plus the uncertainty radius κ multiplied by a constant ($\frac{C}{N}$), the multiplier of s to the left of 1 is negative, and the multiplier of s to the right of 1 is positive, therefore the wealth is entirely allocated to the riskless asset. If the term related to risk of loss on the positive side, $\frac{\gamma^+}{N}$, is dominated by the sum of riskless rate η and the term containing uncertainty radius, then pushing s towards $+\infty$ is the strategy minimizing worst-case risk. In both cases, with increasing riskless rate and uncertainty radius, the investment to the risky market either diminishes, or short positions in the risky market are taken.

For the case with distribution uncertainty based on Kantorovich distance of degree 1, when uncertainty exceeds the uncertainty threshold κ^*, the uniform risky investment strategy is rendered optimal for all possible divisions s and $1-s$ between the riskless and the risky assets. For the Kantorovich distance of degree 2, the allocated amount $1-s$ to risky assets determines the threshold value for the approximate optimality of the uniform investment. Let us assume that the uncertainty radius exceeds a certain level:

$$\kappa \;\geq\; \left[\left(\frac{1}{ND^2}+1\right)^{\frac{1}{2}}+\frac{1}{\sqrt{N}D}\right]\mathbb{E}(\|X^P\|_2^2\mathbb{1}_{\{Z\neq 0\}})^{\frac{1}{2}},$$

for some $D>0$. In this case, the allocation $1-s$ to the risky assets determines the radius of the neighborhood defining the proximity of the optimal allocation of the amount $1-s$ to the the uniform portfolio $w^{u,s}$. The condition (6.56) is satisfied for each $s\neq 0$:

$$\kappa \;\geq\; \left[\left(\frac{(1-s)^2}{ND^2(1-s)^2}+1\right)^{\frac{1}{2}}+\frac{|1-s|}{\sqrt{N}D|1-s|}\right]\mathbb{E}(\|X^P\|_2^2\mathbb{1}_{\{Z\neq 0\}})^{\frac{1}{2}}$$

where the parameter D in Proposition 6.1.3 is replaced with $|1-s|D$, and the optimal solution $w^{*,s}$ of the inner problem (6.5)-(6.6) for the allocation of an amount of $1-s$ to risky assets with fixed s is known to be inside the ball $\{w\in\mathbb{R}^N : \|w-w^{u,s}\|_2<|1-s|D\}$, i.e., at a distance less than $|1-s|D$ to the uniform portfolio $w^{u,s}$ (the case $s=1$, $1-s=0$ is trivial, as discussed above, hence being ignored here). Thus, as the magnitude of the allocation increases, the ability of the uniform risky investment to approximate the optimal allocation reduces.

Let's consider the difference between the objective values (6.5)-(6.6) attains at $w^{*,s}$ and $w^{u,s}$:

$$\begin{aligned}
0 &\leq \mathcal{R}(\langle X^P, w^{u,s}\rangle+\eta s)+C\kappa\|w^{u,s}\|_2-\mathcal{R}(\langle X^P, w^{*,s}\rangle+\eta s)-C\kappa\|w^{*,s}\|_2\\
&\leq C\|w^{u,s}-w^{*,s}\|_2\mathbb{E}(\|X^P\|_2^2\mathbb{1}_{\{Z\neq 0\}})^{\frac{1}{2}}+C\kappa(\|w^{u,s}\|_2-\|w^{*,s}\|_2) \qquad (6.57)\\
&\leq C\|w^{u,s}-w^{*,s}\|_2\mathbb{E}(\|X^P\|_2^2\mathbb{1}_{\{Z\neq 0\}})^{\frac{1}{2}}+C\kappa\|w^{u,s}-w^{*,s}\|_2 \qquad (6.58)\\
&< |1-s|CDE(\|X^P\|_2^2\mathbb{1}_{\{Z\neq 0\}})^{\frac{1}{2}}+|1-s|CD\kappa,
\end{aligned}$$

where the first three inequalities follow due to optimality of $w^{*,s}$, **Lemma 6.1.2**, and the triangle inequality, respectively, and the last inequality is due to the fact that $\|w^{*,s}-w^{u,s}\|_2<|1-s|D$. This means that a band between two functions contains the optimal values for the inner problems (6.5)-(6.6), hence the optimal value for the main problem (6.3)-(6.4) should be sought inside this band:

$$f(s) = \mathcal{R}(\langle X^P, w^{u,s}\rangle + \eta s) + C\kappa\|w^{u,s}\|_2$$
$$g(s) = \mathcal{R}(\langle X^P, w^{u,s}\rangle + \eta s) + C\kappa\|w^{u,s}\|_2 - |1-s|CDE(\|X^P\|_2^2\mathbb{1}_{\{Z\neq 0\}})^{\frac{1}{2}} - |1-s|CD\kappa.$$

When the risk functional choice is $CVaR$, these functions have the following form:

$$\begin{aligned} f(s) &= CVaR_\alpha(-\langle X^P, w^{u,s}\rangle - \eta s) + C\kappa\|w^{u,s}\|_2 \\ &= CVaR_\alpha(-\frac{1-s}{N}\langle \mathbb{1}, X^P\rangle - \eta s) + C\kappa\frac{|1-s|}{\sqrt{N}} \end{aligned}$$

$$\begin{aligned} g(s) &= CVaR_\alpha(-\langle X^P, w^{u,s}\rangle - \eta s) + C\kappa\|w^{u,s}\|_2 - |1-s|CDE(\|X^P\|_2^2\mathbb{1}_{\{Z\neq 0\}})^{\frac{1}{2}} \\ &\quad - |1-s|CD\kappa \\ &= CVaR_\alpha(-\frac{1-s}{N}\langle \mathbb{1}, X^P\rangle - \eta s) + C\kappa\frac{|1-s|}{\sqrt{N}} - |1-s|CDE(\|X^P\|_2^2\mathbb{1}_{\{Z\neq 0\}})^{\frac{1}{2}} \\ &\quad - |1-s|CD\kappa. \end{aligned}$$

Repeating the computations done for the case based on Kantorovich distance of degree 1, and using similar notation, we have:

$$f(s) = \begin{cases} s(-\frac{\gamma^-}{N} - \eta) + \frac{\gamma^-}{N} + C\kappa\frac{|1-s|}{N} & if\ s < 1 \\ \eta & if\ s = 1 \\ s(\frac{\gamma^+}{N} - \eta) - \frac{\gamma^+}{N} + C\kappa\frac{|1-s|}{N} & if\ s > 1, \end{cases}$$

$$g(s) = \begin{cases} s(-\frac{\gamma^-}{N} - \eta) + \frac{\gamma^-}{N} + C\kappa\frac{|1-s|}{N} - |1-s|CD(E(\|X^P\|_2^2\mathbb{1}_{\{Z\neq 0\}})^{\frac{1}{2}} + \kappa) & if\ s < 1 \\ \eta & if\ s = 1 \\ s(\frac{\gamma^+}{N} - \eta) - \frac{\gamma^+}{N} + C\kappa\frac{|1-s|}{N} - |1-s|CD(E(\|X^P\|_2^2\mathbb{1}_{\{Z\neq 0\}})^{\frac{1}{2}} + \kappa) & if\ s > 1. \end{cases}$$

Chapter 7

Computational Experiments

We have a compact set of seven risky assets, avoiding assets causing solution trivialities such as those with very high return and low variance/risk and those with low return and high variance/risk. We compare the single period LPM_1, LPM_2 models with $CVaR$ and VaR models with minimum expected return constraints. We also include as a performance benchmark a mean-variance model, which requires the portfolio to attain an exact level of expected return, while minimizing the return variance:

$$\begin{aligned} \min_{x \in \mathbb{R}^n, s \in \mathbb{R}} \quad & x^T \Gamma x \\ s.t. \quad \mu^T x + Rs &= r_{mv} \\ \mathbb{1}^T x &= 1. \end{aligned}$$

The mean-variance model sets the benchmark attaining a user set expected return level in each setting. Relative performances of the other models in terms of expected return and various risk measures are compared accordingly. For feasibility of the mean-variance model, we make sure that at least two assets with different expected return values exist, so that the target expected return rate r_{mv} is attainable. In this case, one can short-sell the asset with lower return and buy the asset with higher return to attain very high levels of expected return, since there are no lower or upper limits on buying or borrowing/short-selling.

The solution for the risky asset portfolio selection in the mean-variance model is as follows:

$$x^*_{mv} = \frac{r_{mv} - R}{H^2}\Gamma^{-1}\tilde{\mu},$$

where $\tilde{\mu} = \mu - R\mathbb{1}$, and $H^2 = \tilde{\mu}^T\Gamma^{-1}\tilde{\mu}$ is the square of the maximum attainable Sharpe ratio.

There are two settings defining asset first and second moments. In a high-risk setting, the vector defining asset mean returns is as follows:

0.3260	0.2820	0.2573	0.2143	0.1704	0.1393	0.1194

and the matrix defining asset correlations is as follows:

0.7089	0.2380	0.2019	0.2325	0.2157	0.1540	0.1260
0.2380	0.4223	0.2384	0.1636	0.1322	0.1046	0.0990
0.2019	0.2384	0.4218	0.1234	0.1086	0.0526	0.1739
0.2325	0.1636	0.1234	0.3002	0.2262	0.1030	0.0715
0.2157	0.1322	0.1086	0.2262	0.2551	0.0962	0.0475
0.1540	0.1046	0.0526	0.1030	0.0962	0.2233	0.0123
0.1260	0.0990	0.1739	0.0715	0.0475	0.0123	0.1643

Let us consider the first asset in this setting. This asset has a mean return of 0.3260, variance of 0.7089 and a standard deviation of 0.8420. A one standard deviation error in the loss direction results in a return rate of -0.5160, i.e., a 50% loss, and a three standard deviation error in the loss direction results in a 220% loss. One and three standard deviation errors on the lower side result in 5% and 37% losses for the last asset with the lowest variance. Thus, in this setting, we can say that variance dominates the mean return estimate. The maximum attainable Sharpe ratio in this setting is $H = 0.493$.

In the low-risk setting, the mean return vector is as follows:

0.2980	0.2560	0.2333	0.1963	0.1544	0.1253	0.1074

and the asset covariance matrix is as follows:

0.0071	0.0024	0.0020	0.0023	0.0022	0.0015	0.0013
0.0024	0.0042	0.0024	0.0016	0.0013	0.0010	0.0010
0.0020	0.0024	0.0042	0.0012	0.0011	0.0005	0.0017
0.0023	0.0016	0.0012	0.0030	0.0023	0.0010	0.0007
0.0022	0.0013	0.0011	0.0023	0.0026	0.0010	0.0005
0.0015	0.0010	0.0005	0.0010	0.0010	0.0022	0.0001
0.0013	0.0010	0.0017	0.0007	0.0005	0.0001	0.0016

In this setting, the first asset with mean return 0.2980 has a variance of 0.0071 and standard deviation of 0.0842. A one standard deviation error in the loss direction results in a return rate of 0.2138, i.e., a 21% gain, and a three standard deviation error in the loss direction results in a 4% gain. Therefore, in this setting, the effect of variance on the return rate is moderate as a ratio of the mean return rate. The maximum attainable Sharpe ratio in this setting is $H = 4.23$, thus the market conditions are very favorable in contrast to the high-risk setting.

Given the first and second moment information, we test the performances of the models under different distribution types having the same first and second moments. Normal distribution is criticized due to the fact that low tails do not accurately represent historical data. Although we compute performance measures under multivariate Normal asset return distribution, we extend the distribution set with the multivariate Laplace distribution and an affine transformation of the multivariate t-distribution, two distributions from the set of Q-radial distributions, i.e., distributions obtained via affine transformations of elliptical distributions. Both of these distributions have heavier tails than Normal distribution, hence better represent the cases with higher risk of extreme loss. Applicability of elliptical and Q-radial distributions in modeling asset returns is discussed in [71] and [72]. Three multivariate discrete distributions are also considered, where each asset can attain one of two levels of return. In incorporating discrete distributions, the aim is to mimic bound achieving distributions as in Proposition 2.3 of [23]. We have a discrete distribution (Discrete 1), defining assets that attain return rates slightly above the mean with high probability, and far below the mean with very low probability. While having the same mean and covariance as the aforementioned distributions, a second discrete distribution (Discrete 2) defines assets that attain return rates below the mean with high probability, and far above the mean with low probability. In **Figure X**, histogram counts for the return distributions arising from the three continuous distributions under uniform investment (equal distribution of the unit wealth over risky assets) are displayed. The return distribution arising from Normal distribution has the lightest tails, whereas the affine transformation of the t-distribution of degree 2 gives the return distribution with heaviest tails. In **Figure XX**, histogram curves for all five distributions are displayed. Each performance measure is evaluated for each of the five distributions (Normal, Laplace, (transformed) t-distribution of degree 2, Discrete 1 and Discrete 2 distributions). The worst-case performance over all five distribution types is computed as the lowest return level or highest risk measure attained throughout all distribution types. A fictive average-case for each performance measure is obtained via taking averages over the performance measure evaluations for the five distribution types. Computation of the worst-case behavior under ellipsoidal mean return uncertainty is done based on the theoretical bound (a loss of $-\epsilon\sqrt{x^T\Gamma^{-1}x}$ in returns, where x is the solution of the relevant model) rather than generating uniform ellipsoidal perturbations and computing the worst-case performance. The former method is more accurate, and computationally less expensive.

Before comparing the performances of the solutions of models under different settings, let us remind the solutions of the other three models to be compared, $x^*_{LPM_1}$, $x^*_{LPM_2}$ and x^*_{CVaR}, i.e., the worst-case LPM_1, LPM_2 and $CVaR$ minimizing portfolio under the existence of a riskless asset, distribution ambiguity and mean return vector uncertainty:

$$x^*_{LPM_1} = 2(r - R)\frac{H - \epsilon}{H + H(H - \epsilon)^2}\Gamma^{-1}\tilde{\mu} \ ,$$

$$x^*_{LPM_2} = (r - R)\frac{(H + \epsilon)(H^2 - \epsilon^2)}{H[(H + \epsilon)^2 + (H^2 - \epsilon^2)^2]}\Gamma^{-1}\tilde{\mu} \ ,$$

$$x^*_{CVaR} = (d - R)\frac{1}{(H - \epsilon)H}\Gamma^{-1}\tilde{\mu} \ .$$

All of the solutions above are actually the vector $\Gamma^{-1}\tilde{\mu}$, in the direction maximizing Sharpe ratio, scaled by coefficients specific to the model:

$$c_{mv}(\epsilon) = (r_{mv} - R)\frac{1}{H^2} \ ,$$

$$c_{LPM_1}(\epsilon) = 2(r - R)\frac{H - \epsilon}{H + H(H - \epsilon)^2} \ ,$$

$$c_{LPM_2}(\epsilon) = (r - R)\frac{(H + \epsilon)(H^2 - \epsilon^2)}{H[(H + \epsilon)^2 + (H^2 - \epsilon^2)^2]} \ ,$$

$$c_{CVaR}(\epsilon) = (d - R)\frac{1}{(H - \epsilon)H} \ .$$

These coefficients have some similarities, for instance, in the limit (as $\epsilon \to 0$ or $H \to \infty$) the dominant term is $\frac{1}{H^2}$. H, the maximum attainable Sharpe ratio, indicates how favorable the market conditions are. A higher value for H means securing investment with a certain expected return at a lower level of risk/variance. For fixed Γ, for instance, if H increases due to doubling $\tilde{\mu}$, the risky investment is (approximately) cut down in half, that is, optimality is attained with half the risky investment. Again, models are similar in that the coefficients increase with increasing target rates of models (r_{mv}, r, d), and decrease with increasing riskless rate R. hence increasing target rates force the investor into the risky market, whereas increasing riskless rate does the opposite. The multiplier 2 in c_{LPM_1}

should also be noted, this reflects the moderate risk aversion in comparison to the LPM_2-minimizing model. The comparison of models regarding the change in anticipated mean-return uncertainty radius, ϵ, should be easier based on graphic illustration. In **Figure 7.1**, $c_{CVaR}(\epsilon)$ explodes, since for the $CVaR$ model, the nominator is constant with respect to ϵ while the term $H - \epsilon$ drives the denominator to zero as ϵ approaches H. This is due to the expected return constraint in the $CVaR$ model. The worst-case attainable Sharpe ratio approaches zero as ϵ converges to H, and the investor has to increase the investment amounts to satisfy the expected return constraint. Thus the investment amount approaches infinity as ϵ approaches H. On the other hand, c_{LPM_1} and c_{LPM_2} have terms that keep the denominator bounded away from zero in the positive side, while a multiplicative term drives nominators to zero as ϵ approaches H. In **Figure 7.2**, c_{CVaR} is removed from the picture, for the behavior of c_{LPM_1} and c_{LPM_2} to be observed clearly. The term form risky assets approaches zero for both models as ϵ approaches H. In the figures presented in this chapter, the x-axis is a dummy axis for the mean-variance model, since this model does not incorporate an uncertainty anticipation for the mean return vector estimation. Thus, while the performances for the other models change with respect to the mean return uncertainty incorporated in the model, performance for the mean-variance model is represented by a flat line in each case. In the figures, the term $Model\,\epsilon$ is used to refer to the level of uncertainty the investor anticipates and incorporates in the model; hence, this is the ϵ appearing in the above mentioned optimal portfolio selections. On the other hand, in each setting, i.e., for each figure, the actual radius of mean return uncertainty is a constant, referred as the $Market\,\epsilon$. This amount is used as a basis for the perturbations in generated market return data and investment returns. Considering performance for the cases when $Model\,\epsilon$ is below and above $Market\,\epsilon$, it is possible to observe results of under- and overestimation of the level of uncertainty in the mean return vector estimate.

Note the selections for the parameters in **Figures 7.1 and 7.2**. The parameter r_{mv} in the mean-variance model sets an exact level of expected return for the portfolio; the portfolio selection exactly attains this level. We use this parameter to set a benchmark line in the plots for expected return on portfolio.

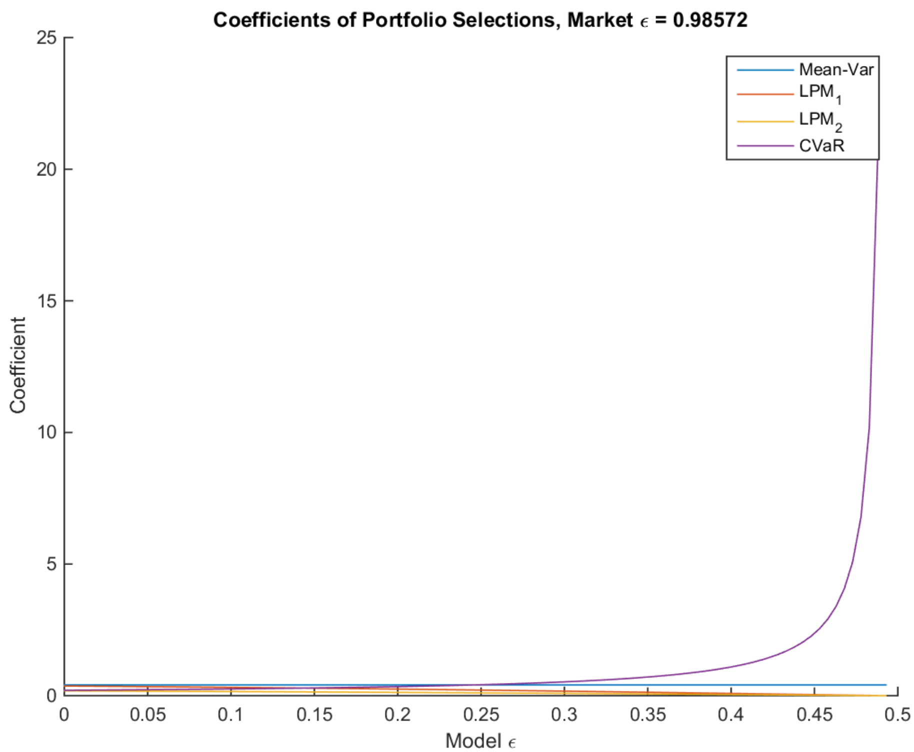

Figure 7.1: Change in Coefficients as Anticipated Mean Uncertainty Radius Increases. $H = 0.493$, $r_{mv} = 0.12$, $r = 0.25$, $d = 0.07$.

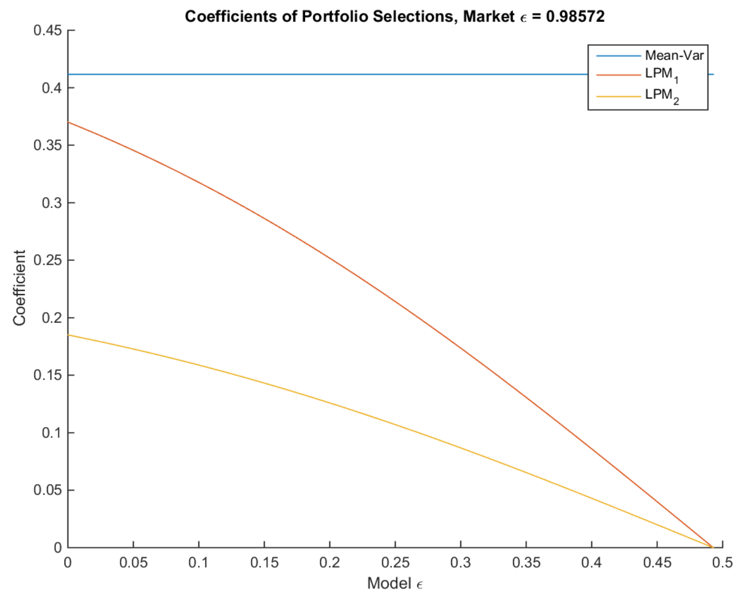

Figure 7.2: Change in Coefficients as Anticipated Mean Uncertainty Radius Increases, without the $CVaR$ model. $H = 0.493$, $r_{mv} = 0.12$, $r = 0.25$, $d = 0.07$.

The parameter r in the LPM_1 and LPM_2 model solutions, on the other hand, is in general selected to be a large value, indicating a "utopic/optimistic" target. Especially in a high-risk setting, the expected return rate attained by the optimal portfolio selections of the LPM_1 and LPM_2 models fall below the target rate r. Although investor preferences might lead to different choices on the target rate in the LPM_1 and LPM_2 models, we set a single target rate r for both LPM_1 and LPM_2 models, to be able to compare the behavior under different levels of risk aversion, as indicated by the degree of the lower partial moment, $m = 1, 2$. The parameter d in the $CVaR$ model solution sets a lower bound for the expected investment return under the worst-case realization of the mean return vector. Therefore, it is sensible to set a value for d below that of r_{mv}, and it is sensible to set a value for r above that of r_{mv}, especially in a high-risk setting.

In **Figures 7.3, 7.4, 7.5, 7.6**, the portfolio selection profiles of all four models in a high-risk setting can be observed. As noted above, the risky asset selections are vectors in the same direction for all models. Therefore, the same risky assets are bought and short-sold with the same ratio in all models. The selection for

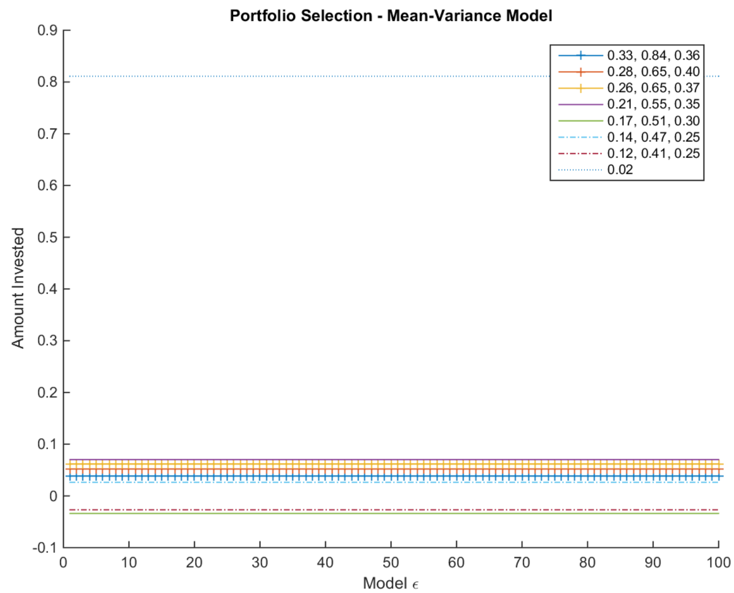

Figure 7.3: Portfolio Selection for the Mean-Variance Model. High-risk Setting. Constant over x-axis since the model does not involve mean return vector uncertainty. Legend shows return mean, standard deviation and Sharpe-ratio for each single risky asset; return rate for the riskless asset. $H = 0.493$, $r_{mv} = 0.12$, $r = 0.25$, $d = 0.07$.

LPM_1 model has twice the allocation for risky assets compared to the LPM_2 model, and the allocation to the riskless asset is therefore higher for the LPM_2 model. In the high-risk setting, as the investor anticipates higher level of uncertainty in the mean return vector, the investment into the risky assets diminishes, if the LPM models are used. For the $CVaR$ model, the opposite behavior is observed, as the mean return uncertainty incorporated into the model increases, the investor scales up the risky investment to satisfy the lower bound constraint on expected portfolio return. In the low-risk environment, both LPM models and the $CVaR$ model display the same behavior: as the investor anticipates a higher level of uncertainty in the mean return vector, the investor reacts by scaling up the risky investment (**Figures 7.7, 7.8, 7.9**).

In the high risk environment, over-estimation of the level of mean return vector uncertainty results in reduced risks at the cost of lowered gains, that is, expected portfolio return rate (**Figures 7.10, 7.11, and 7.12**). This being true

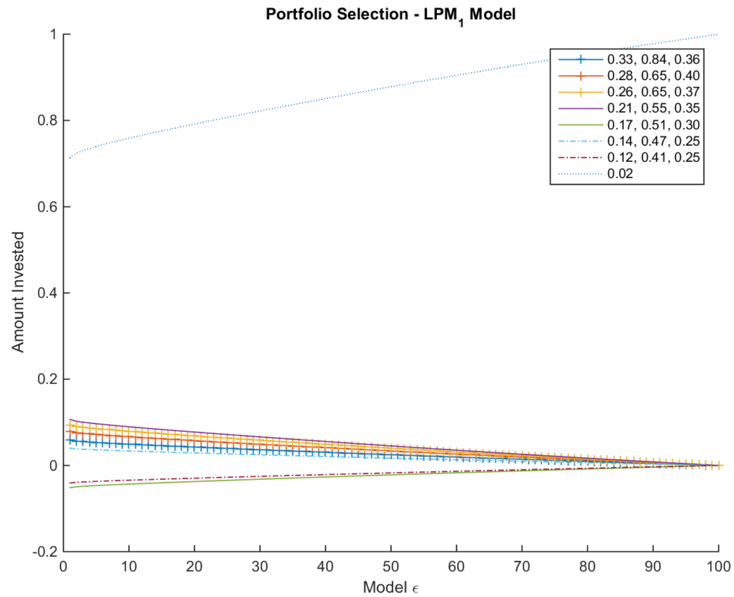

Figure 7.4: Portfolio Selection for the LPM_1 Model. High-risk Setting. Legend shows return mean, standard deviation and Sharpe-ratio for each single risky asset; return rate for the riskless asset. $H = 0.493$, $r_{mv} = 0.12$, $r = 0.25$, $d = 0.07$.

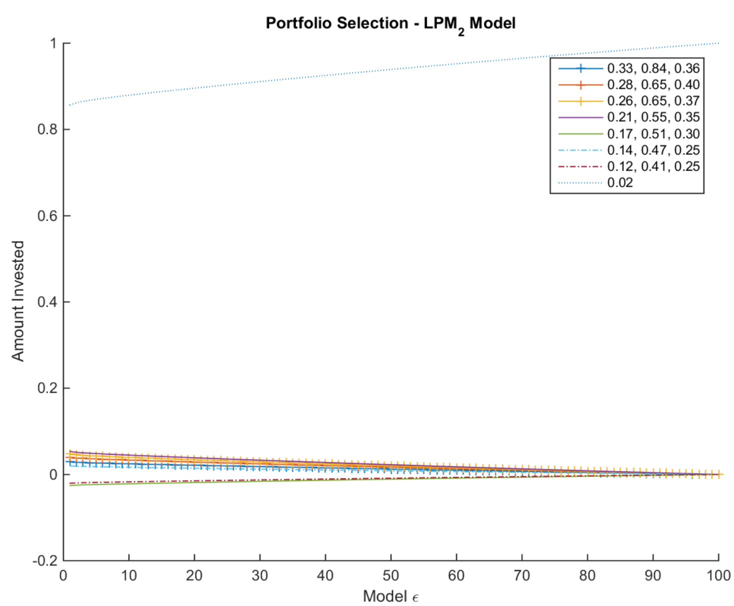

Figure 7.5: Portfolio Selection for the LPM_2 Model. High-risk Setting. Legend shows return mean, standard deviation and Sharpe-ratio for each single risky asset; return rate for the riskless asset. $H = 0.493$, $r_{mv} = 0.12$, $r = 0.25$, $d = 0.07$.

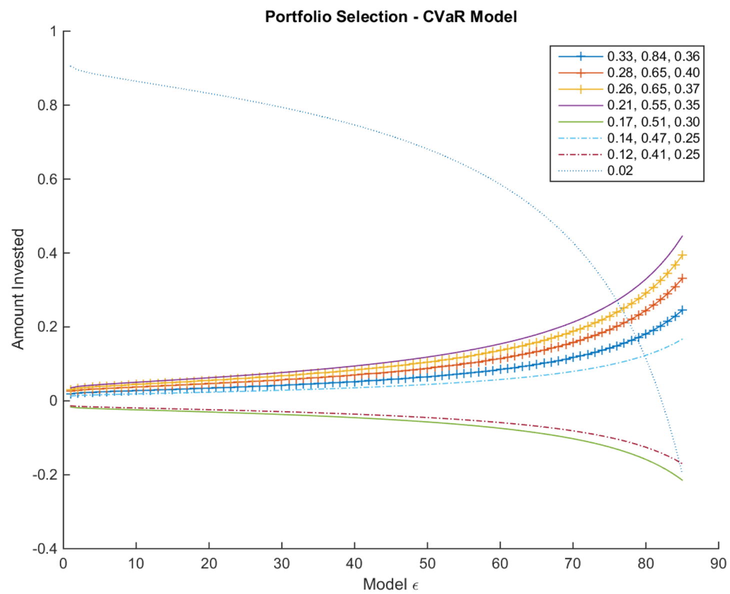

Figure 7.6: Portfolio Selection for the $CVaR$ Model. High-risk Setting. Legend shows return mean, standard deviation and Sharpe-ratio for each single risky asset; return rate for the riskless asset. $H = 0.493$, $r_{mv} = 0.12$, $r = 0.25$, $d = 0.07$.

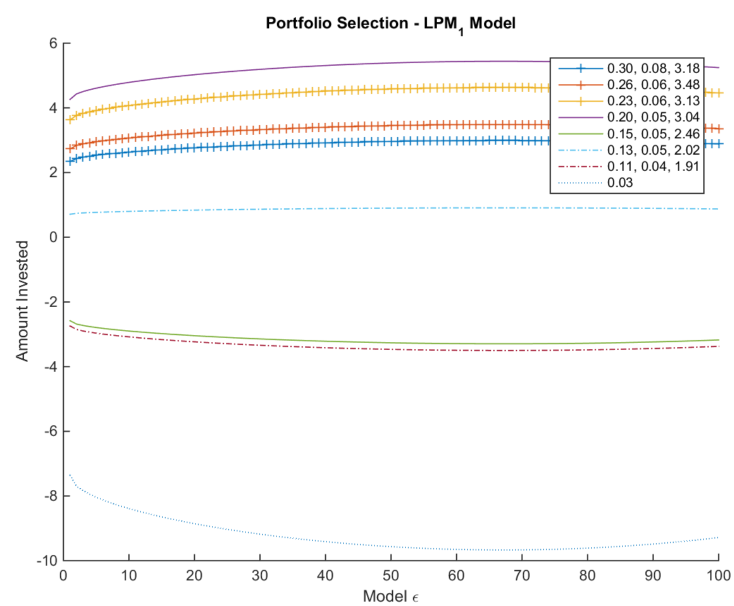

Figure 7.7: Portfolio Selection for the LPM_1 Model. Low-risk Setting. Legend shows return mean, standard deviation and Sharpe-ratio for each single risky asset; return rate for the riskless asset. $H = 4.29$, $r_{mv} = 0.45$, $r = 0.35$, $d = 0.40$.

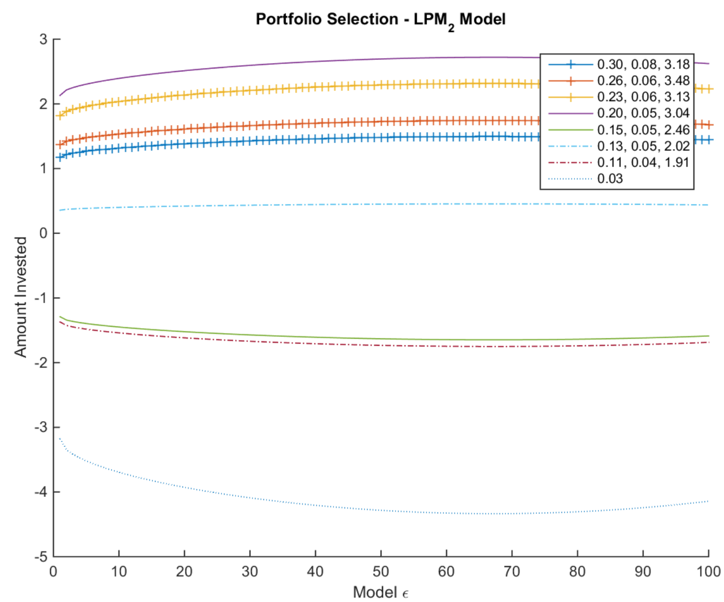

Figure 7.8: Portfolio Selection for the LPM_2 Model. Low-risk Setting. Legend shows return mean, standard deviation and Sharpe-ratio for each single risky asset; return rate for the riskless asset. $H = 4.29$, $r_{mv} = 0.45$, $r = 0.35$, $d = 0.40$.

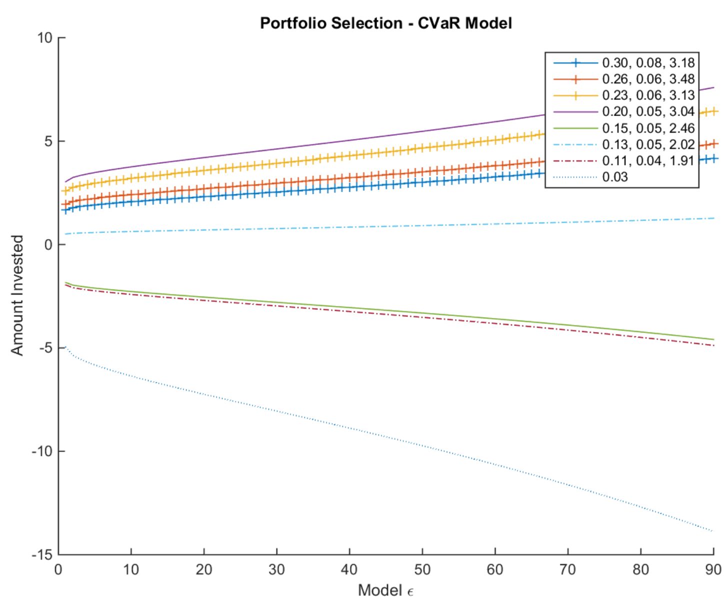

Figure 7.9: Portfolio Selection for the $CVaR$ Model. Low-risk Setting. Legend shows return mean, standard deviation and Sharpe-ratio for each single risky asset; return rate for the riskless asset. $H = 4.29$, $r_{mv} = 0.45$, $r = 0.35$, $d = 0.40$.

for the LPM models, for the $CVaR$ model, overestimation of ϵ results in growing risky investment, increased risk measures and increased expected portfolio return. Since the $CVaR$ model is forced to scale up risky investment on the anticipation of high mean return vector uncertainty, worst-case $CVaR$ risk measure also grows for this model (**Figure 7.13**). However, gains on an average-case do not necessarily imply gains in the worst-case. In **Figure 7.10**, it can be seen that on an average-case, expected portfolio return increases as the risky investment is scaled up in the $CVaR$ model. However, worst-case behavior is presented in **Figure 7.14**: the attempt of the $CVaR$ model investor to secure the mean return constraint results in a disaster in terms of worst-case expected portfolio return. For the LPM models, the opposite is observed, conservatism increases with model uncertainty radius, reducing portfolio returns on an average-case, but actually the worst-case expected portfolio return increases with increasing uncertainty radius incorporated in the model. Similarly, LPM_1 and LPM_2 risk measures increase as model ϵ increases, on an average-case, but actually the worst-case measures are reduced (compare **Figures 7.11 and 7.12** with **Figures 7.15 and 7.16**). Note that the models discussed in this study aim to minimize worst-case risk under distribution ambiguity and mean return vector uncertainty, rather than minimizing an average-case risk or risk measures for a specific distribution type. Therefore, on the average, the risk-measures for the models might be higher than that of the mean-variance model, or increase with a correct anticipation of mean return vector uncertainty. However, for the actual focus of interest of the models, i.e., the worst-case risk measures, performance is better than the mean-variance model, and incorporating mean return vector uncertainty in the model further reduces risk measures.

Overestimation of the mean return vector uncertainty can have complex effects in terms of risk measures. In a low-risk setting, if the investor anticipates a high level of mean return vector uncertainty, he responds by increasing the risky investment. Due to the low risk level, this helps increasing expected portfolio return even in the worst-case (**Figure 7.17**). This results in reduced worst-case risk for some measures (**Figures 7.18 and 7.19**), while increasing others (**Figure 7.20**). In **Figure 7.20**, the success due to accurate estimation of the mean return

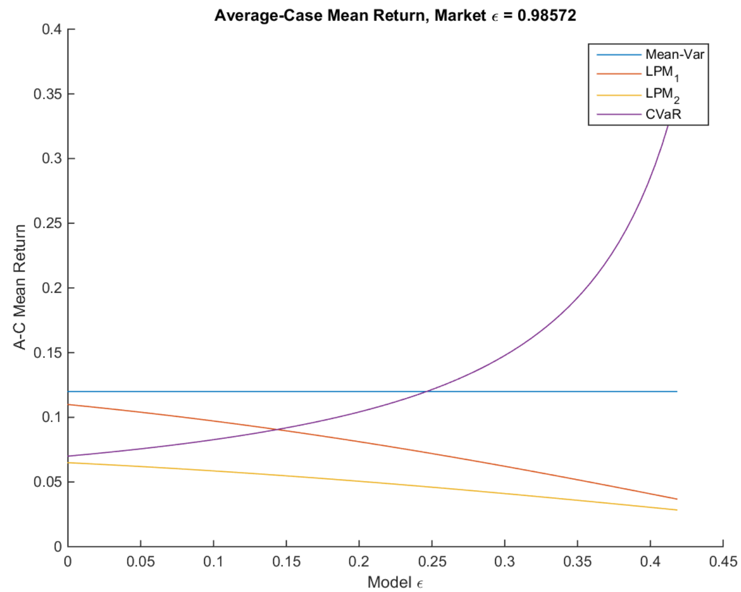

Figure 7.10: Expected Return Rate on Portfolio on an Average-Case. High-risk Setting. $H = 0.493$, $r_{mv} = 0.12$, $r = 0.25$, $d = 0.07$.

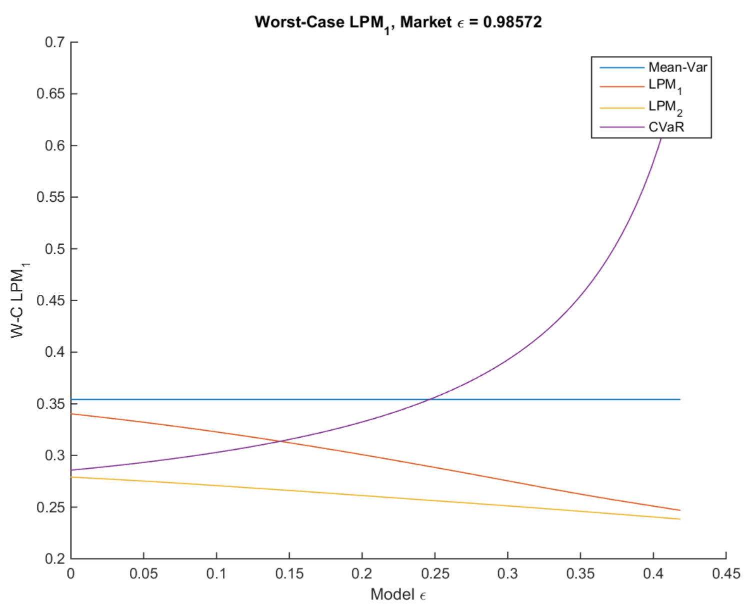

Figure 7.11: Worst-case LPM_1 measures. High-risk Setting. $H = 0.493$, $r_{mv} = 0.12$, $r = 0.25$, $d = 0.07$.

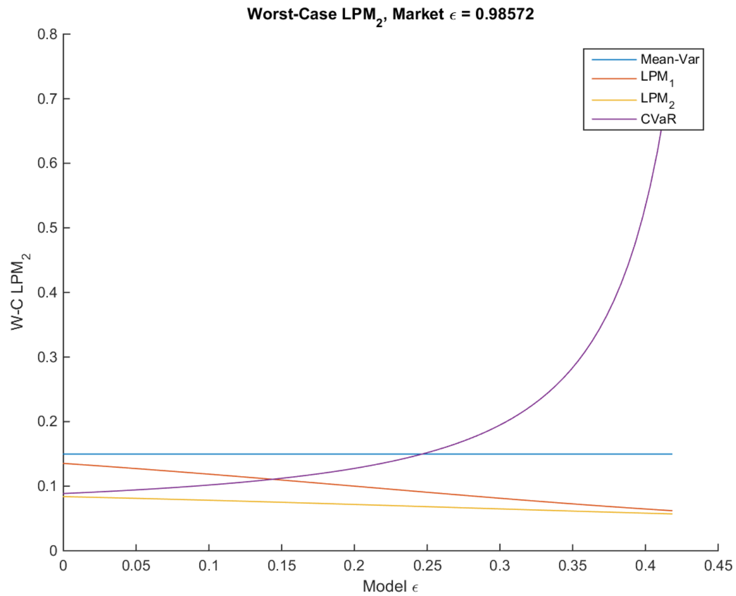

Figure 7.12: Worst-case LPM_2 measures. High-risk Setting. $H = 0.493$, $r_{mv} = 0.12$, $r = 0.25$, $d = 0.07$.

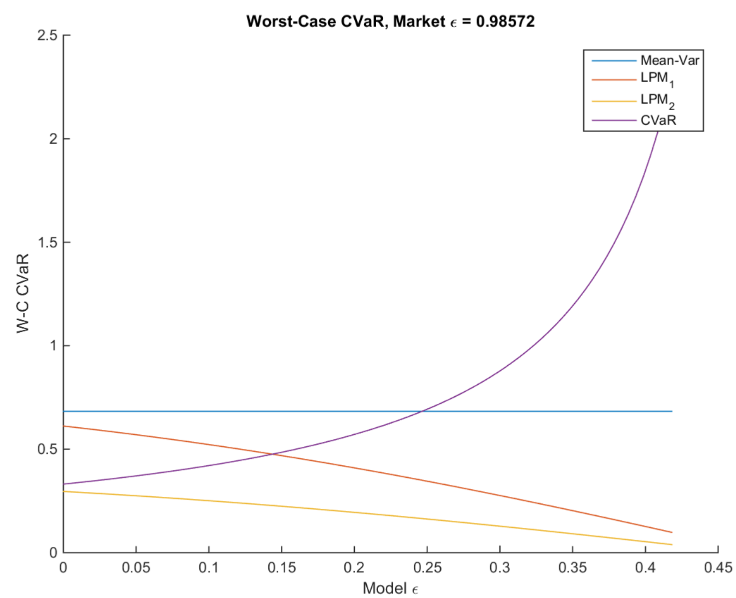

Figure 7.13: Worst-case $CVaR$ measures. High-risk Setting. $H = 0.493$, $r_{mv} = 0.12$, $r = 0.25$, $d = 0.07$.

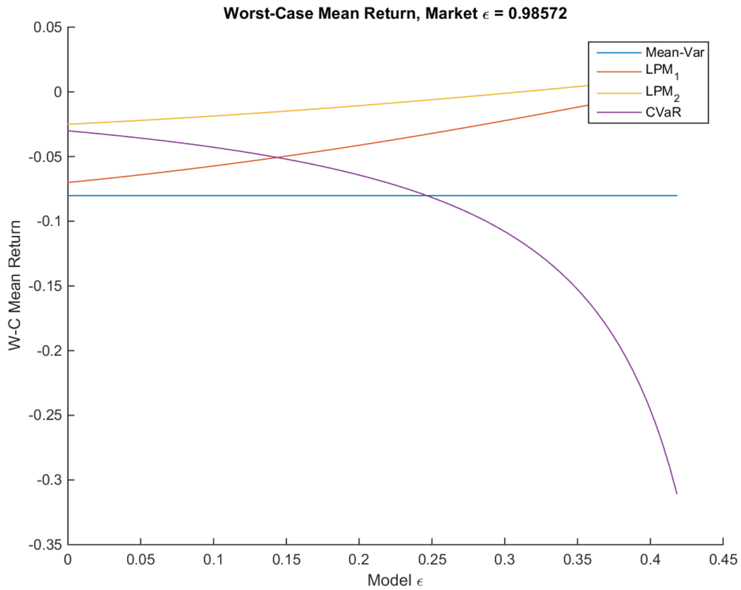

Figure 7.14: Worst-case $CVaR$ measures. High-risk Setting. $H = 0.493$, $r_{mv} = 0.12$, $r = 0.25$, $d = 0.07$.

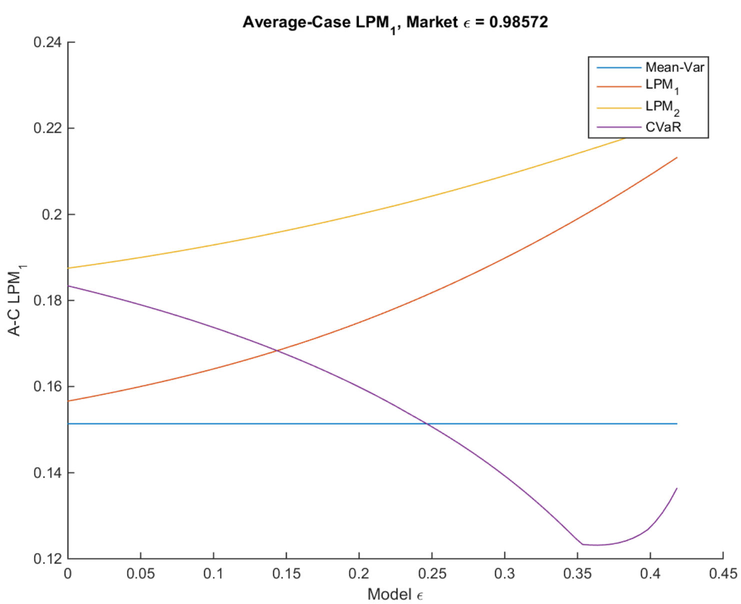

Figure 7.15: Average-case LPM_1 measures. High-risk Setting. $H = 0.493$, $r_{mv} = 0.12$, $r = 0.25$, $d = 0.07$.

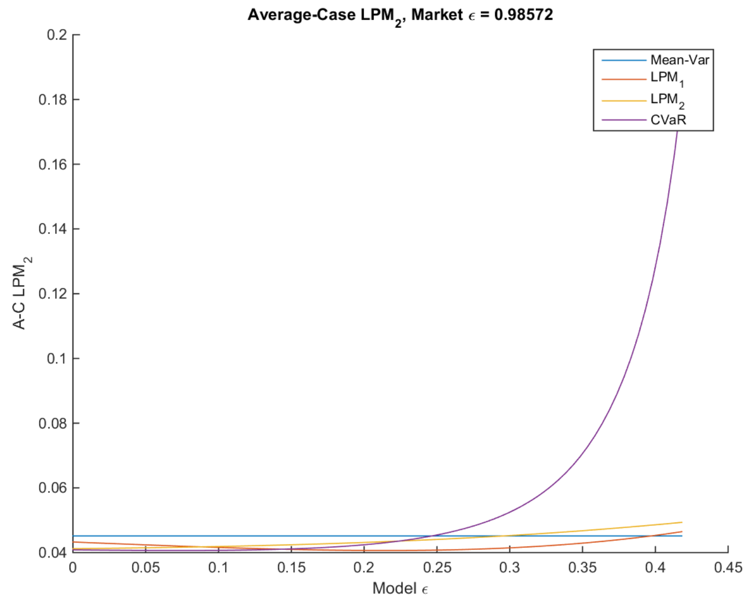

Figure 7.16: Average-case LPM_2 measures. High-risk Setting. $H = 0.493$, $r_{mv} = 0.12$, $r = 0.25$, $d = 0.07$.

uncertainty level in the market can be observed. While the worst-case LPM_2 risk measure increases for LPM_1 and $CVaR$ model solutions as higher mean return vector uncertainty is incorporated in the models, the minimum for this measure is attained with an accurate modelling of ϵ for the LPM_2 model.

Shifting the distribution right without causing significant spreading effectively reduces the (worst-case) LPM_1 measure. In **Figure 7.21**, we observe that the LPM_1 model achieves this even when no mean return uncertainty is involved. LPM_2 and $CVaR$ models on the other hand apply this shift with increasing uncertainty radius, trying to reduce their corresponding risk measures, on the mean time reducing the worst-case LPM_1 risk measure.

It is not surprising to see that the worst-case performance curves are mostly defined by the first discrete distribution, which gives a highly left-skewed return distribution with low probability extreme losses. In the high-risk setting, performance for different distribution types are defined by more or less similar curves (see, for instance, **Figure 7.22** or **Figure 7.23**). In the low-risk setting, distribution type turns to be more important, where curves for the first discrete

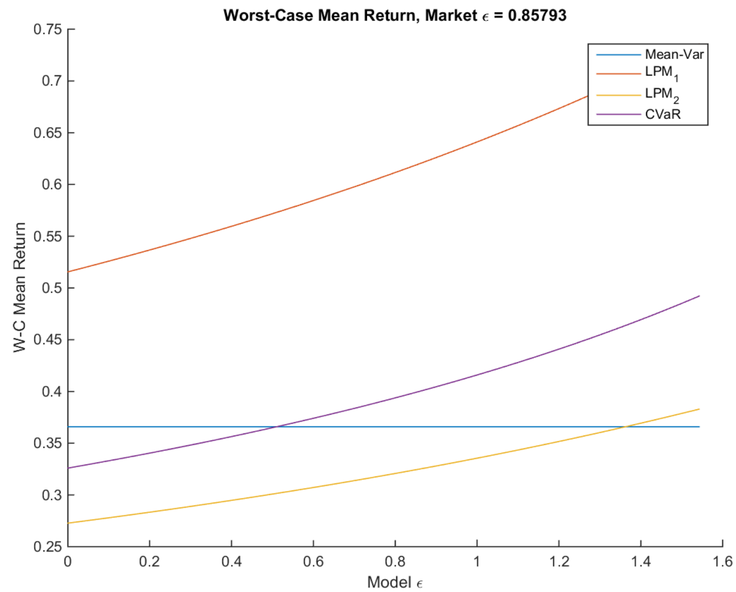

Figure 7.17: Worst-case Expected Portfolio Return. Low-risk Setting. $H = 4.29$, $r_{mv} = 0.45$, $r = 0.35$, $d = 0.40$.

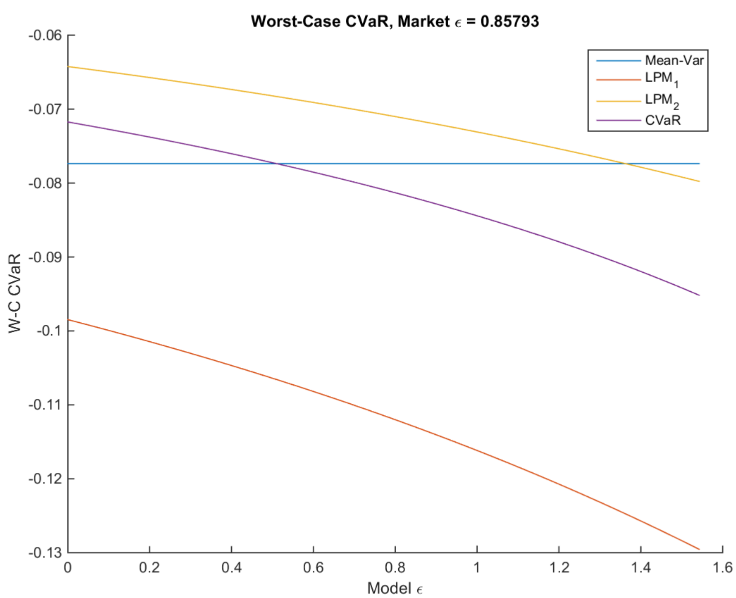

Figure 7.18: Worst-case $CVaR$ measures. Low-risk Setting. $H = 4.29$, $r_{mv} = 0.45$, $r = 0.35$, $d = 0.40$.

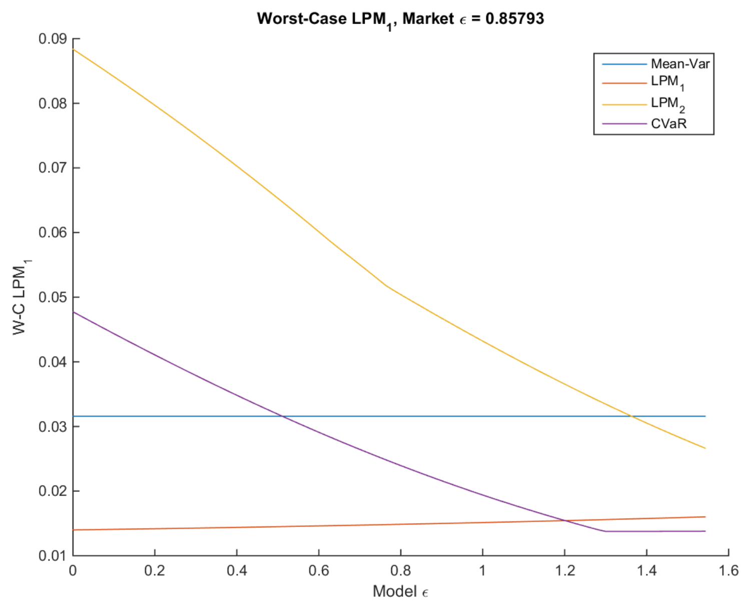

Figure 7.19: Worst-case LPM_1 measures. Low-risk Setting. $H = 4.29$, $r_{mv} = 0.45$, $r = 0.35$, $d = 0.40$.

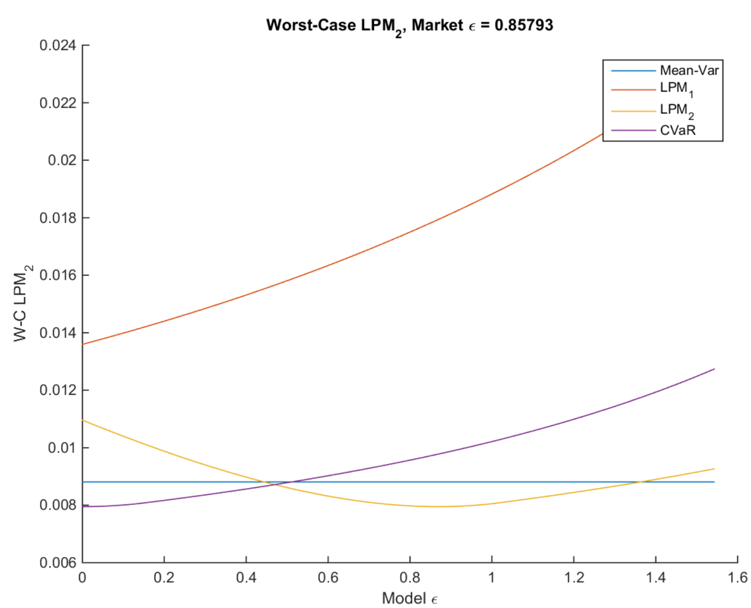

Figure 7.20: Worst-case LPM_2 measures. Low-risk Setting. $H = 4.29$, $r_{mv} = 0.45$, $r = 0.35$, $d = 0.40$.

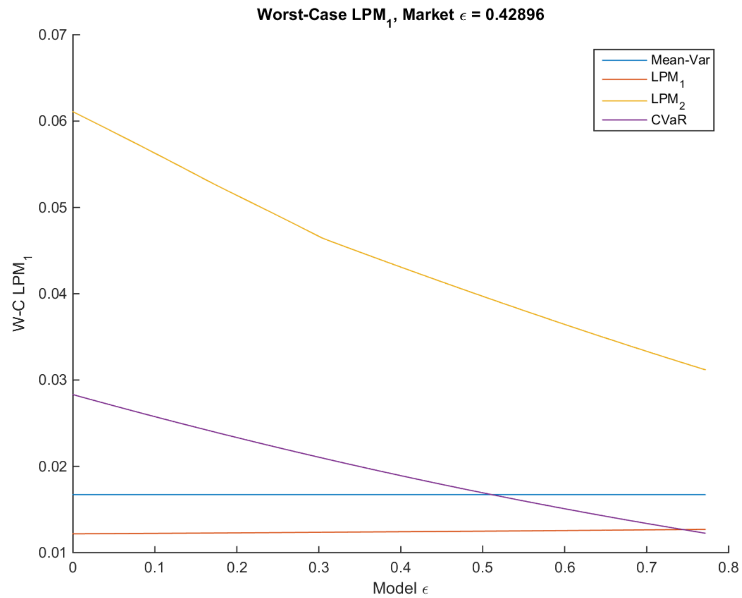

Figure 7.21: Worst-case LPM_1 measures. Low-risk Setting. $H = 4.29$, $r_{mv} = 0.45$, $r = 0.35$, $d = 0.40$.

distribution are far off from those for other distributions (see, for instance, **Figure 7.24** or **Figure 7.25**).

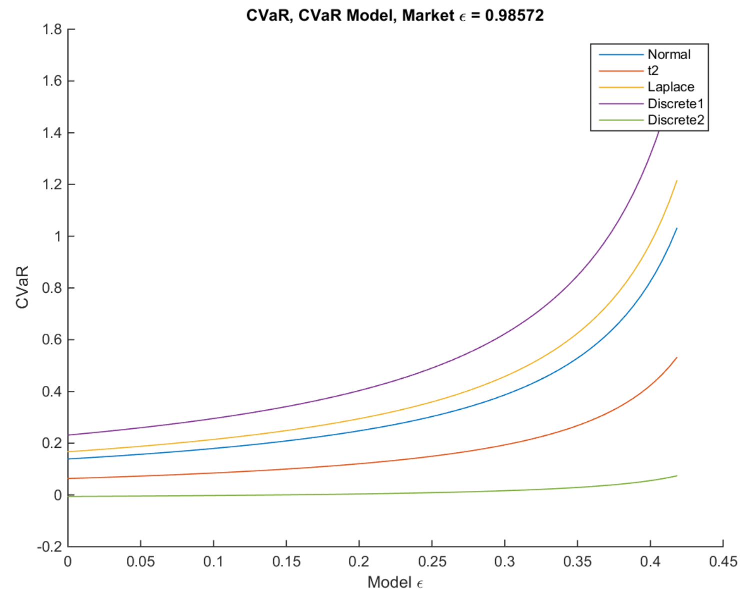

Figure 7.22: $CVaR$ Risk Measures for Return Distributions Based on $CVaR$ Model Solution and Different Distribution Types. High-risk Setting. $H = 0.493$, $r_{mv} = 0.12$, $r = 0.25$, $d = 0.07$.

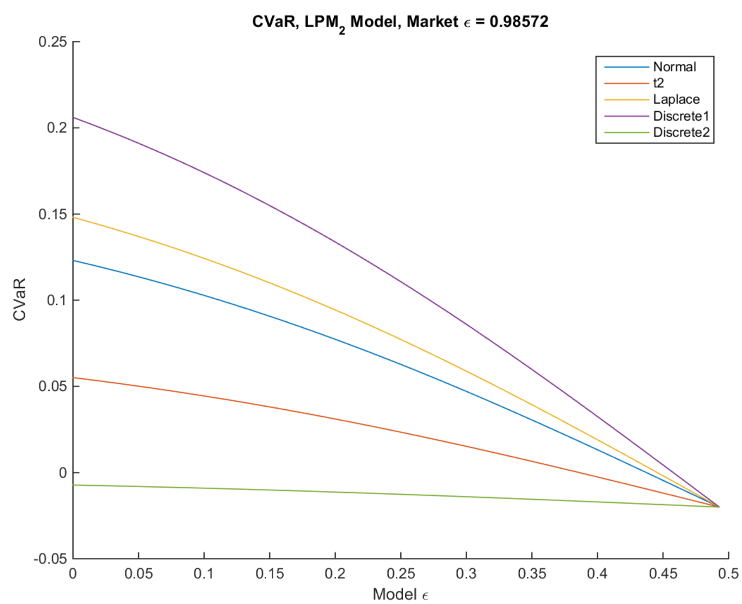

Figure 7.23: $CVaR$ Risk Measures for Return Distributions Based on LPM_2 Model Solution and Different Distribution Types. High-risk Setting. $H = 0.493$, $r_{mv} = 0.12$, $r = 0.25$, $d = 0.07$.

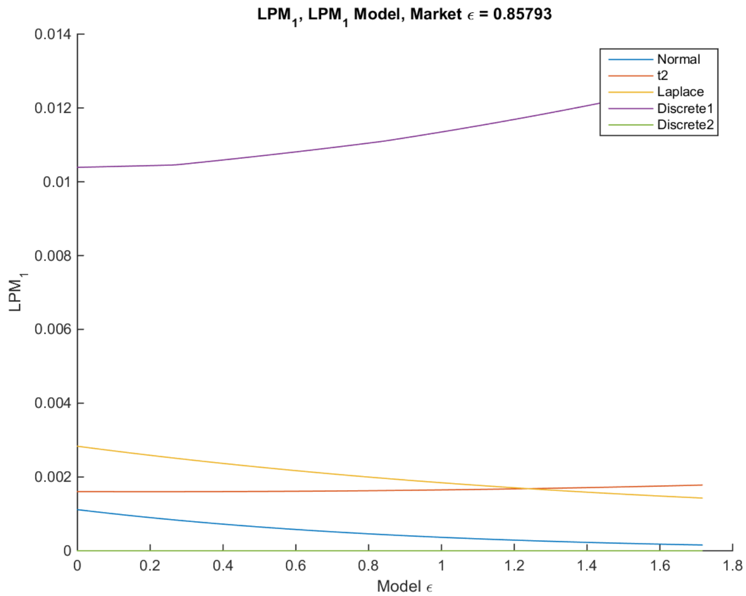

Figure 7.24: LPM_1 Risk Measures for Return Distributions Based on LPM_1 Model Solution and Different Distribution Types. Low-risk Setting. $H = 4.29$, $r_{mv} = 0.45$, $r = 0.35$, $d = 0.40$.

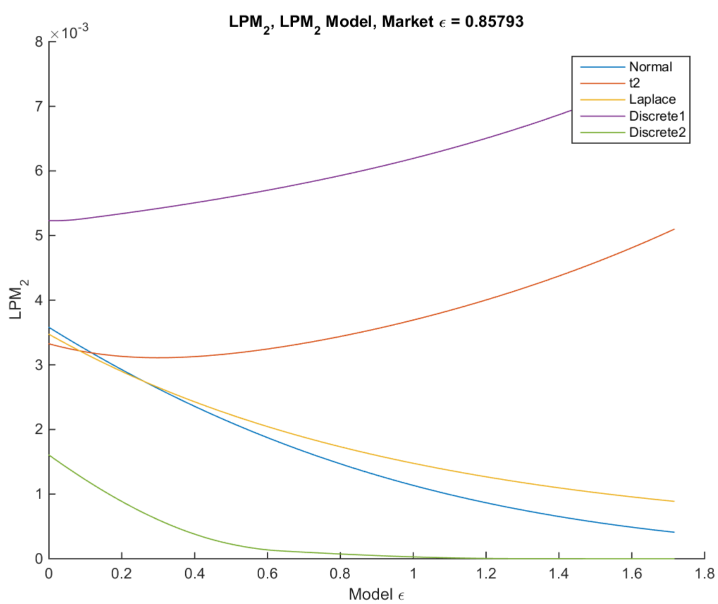

Figure 7.25: LPM_2 Risk Measures for Return Distributions Based on LPM_2 Model Solution and Different Distribution Types. Low-risk Setting. $H = 4.29$, $r_{mv} = 0.45$, $r = 0.35$, $d = 0.40$.

Chapter 8

Conclusion

In this study, we applied robust modeling and robust optimization approach on portfolio selection. We studied the risk minimizing portfolio selection problem under different risk measures and different perspectives regarding distributional uncertainty. In the first three chapters, we restricted the information on the distribution of asset returns to the first two moments, minimizing the worst-case risk measure over all possible distributions whose first two moments match the information at hand. For the majority of the problems studied here, the investment environment involves several risky assets and a riskless asset. Short positions are allowed for both risky and riskless assets. In Chapters 3 and 4, we minimized the (worst-case) first and second lower partial moment risk measures, considering distribution uncertainty cases and extending these with ambiguity in the mean return. We derived closed-form portfolio under these conditions for both static and dynamic problems. The behavior of the risk measures when combined with the worst-case/supremum operator rendered myopic single period solutions optimal for dynamic settings. In Chapter 5, we considered worst-case CVaR and worst-case VaR measures. The investment environment is defined in a similar manner to those in Chapters 3 and 4, as well as the distributional uncertainty and the ellipsoidal mean return uncertainty extension. With a minimum expected portfolio return constraint, we incorporated in the model the investor's expectation for gains. Although this constraint made the models in Chapter 5 comparable with the models in the earlier chapters, the mean-variance efficient optimal portfolio rules indicate aggressive behavior regarding risk, as a response to increasing uncertainty in the mean return vector. While the results in Chapters 3 and 4 advise the investor to deal with increasing mean vector uncertainty by shifting allocation to the riskless asset, the results in Chapter 5 advise borrowing cash and investing into risky assets as a response to increasing mean vector uncertainty.

In a final problem, we have a change of perspective regarding uncertainty, and focus our attention on naive diversification strategies. Prevalence of naive diversification strategies influencing choice in many contexts is an intriguing phenomenon for both behavioral science and financial studies. While behavioral studies investigate the psychological basis of this choice that is seemingly irrational in contradicting theoretically optimal decisions, in the context of portfolio selection, naive diversification is considered as a plausible reaction against uncertainty. A mathematical model justifying the rationale behind naive diversification is very valuable, explaining the prevalence of this behavior. A previous result on the optimality of uniform portfolio under increasing uncertainty is a solid evidence supporting the rationality of naive diversification strategies. In Chapter 6, we adopted this framework modeling worst-case risk minimization under uncertainty. Here, the knowledge of a nominal multi-variate distribution of asset returns is assumed. The actual distribution is considered to be within a ball around this nominal distribution, defined by the Kantorovich distance. The choice of the metric in this model is particularly due to its weak convergence properties, since under this metric it is possible to show that with increasing uncertainty radius, the optimal worst-case risk minimizing portfolio converges to the uniform portfolio - the naive diversification strategy. Adopting this framework, we extend the model to investigate convergence of the risky investment to uniform portfolio under the existence of a riskless asset. Exploiting the riskless asset, we intend to elicit the attitude of the investor towards taking risk in an environment where naive diversification is justified to be an optimal strategy. Indeed, while it renders uniform investment to risky assets optimal, increasing uncertainty radius causes diminishing allocation to risky assets. Hence, as uncertainty increases, the risk averse investor is driven out of the risky market.

Bibliography

[1] J. M. Mulvey, R. J. Vanderbei, and S. A. Zenios, "Robust optimization of large-scale systems," *Operations Research*, vol. 43, no. 2, pp. 264–281, 1995.

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

[3] S. Benati and R. Rizzi, "A mixed integer linear programming formulation of the optimal mean/value-at-risk portfolio problem," *European Journal of Operational Research*, vol. 176, no. 1, pp. 423–434, 2007.

[4] P. Artzner, F. Delbaen, J.-M. Eber, and D. Heath, "Coherent measures of risk," *Mathematical Finance*, vol. 9, no. 3, pp. 203–228, 1999.

[5] G. C. Pflug, A. Pichler, and D. Wozabal, "The $1/n$ investment strategy is optimal under high model ambiguity," *Journal of Banking & Finance*, vol. 36, no. 2, pp. 410–417, 2012.

[6] A. Ben-Tal and A. Nemirovski, "Robust solutions of uncertain linear programs," *Operations Research Letters*, vol. 25, no. 1, pp. 1–13, 1999.

[7] A. Ben-Tal, A. Nemirovski, and C. Roos, "Robust solutions of uncertain quadratic and conic-quadratic problems," *SIAM Journal on Optimization*, vol. 13, no. 2, pp. 535–560, 2002.

[8] A. Ben-Tal, A. Goryashko, E. Guslitzer, and A. Nemirovski, "Adjustable robust solutions of uncertain linear programs," *Mathematical Programming*, vol. 99, no. 2, pp. 351–376, 2004.

[9] L. El Ghaoui and H. Lebret, "Robust solutions to least-squares problems with uncertain data," *SIAM Journal on Matrix Analysis and Applications*, vol. 18, no. 4, pp. 1035–1064, 1997.

[10] L. El Ghaoui, F. Oustry, and H. Lebret, "Robust solutions to uncertain semidefinite programs," *SIAM Journal on Optimization*, vol. 9, no. 1, pp. 33–52, 1998.

[11] L. Vandenberghe and S. Boyd, "Semidefinite programming," *SIAM Review*, vol. 38, no. 1, pp. 49–95, 1996.

[12] A. Nemirovski and Y. Nesterov, "Interior-point polynomial methods in convex programming," *Lecture notes, Faculty of Industrial Engineering and Management, Israel Institute of Technology, Israel*, 1994.

[13] D. Bertsimas, D. Pachamanova, and M. Sim, "Robust linear optimization under general norms," *Operations Research Letters*, vol. 32, no. 6, pp. 510–516, 2004.

[14] A. Ben-Tal and A. Nemirovski, "On tractable approximations of uncertain linear matrix inequalities affected by interval uncertainty," *SIAM Journal on Optimization*, vol. 12, no. 3, pp. 811–833, 2002.

[15] D. Bertsimas and M. Sim, "Tractable approximations to robust conic optimization problems," *Mathematical Programming*, vol. 107, no. 1-2, pp. 5–36, 2006.

[16] A. Ben-Tal and A. Nemirovski, "On approximate robust counterparts of uncertain semidefinite and conic quadratic programs," in *System Modeling and Optimization XX*, pp. 1–22, Springer, 2003.

[17] A. Ben-Tal and A. Nemirovski, "Robust optimization–methodology and applications," *Mathematical Programming*, vol. 92, no. 3, pp. 453–480, 2002.

[18] D. Bertsimas, D. B. Brown, and C. Caramanis, "Theory and applications of robust optimization," *SIAM review*, vol. 53, no. 3, pp. 464–501, 2011.

[19] H. M. Markowitz, *Portfolio selection: efficient diversification of investments*, vol. 16. Yale University Press, 1968.

[20] R. C. Merton *et al.*, "An analytic derivation of the efficient portfolio frontier," *Journal of Financial and Quantitative Analysis*, vol. 7, no. 4, pp. 1851–1872, 1972.

[21] R. Merton, "Continuous-time finance," 1990.

[22] M. C. Steinbach, "Markowitz revisited: Mean-variance models in financial portfolio analysis," *SIAM Review*, vol. 43, no. 1, pp. 31–85, 2001.

[23] L. Chen, S. He, and S. Zhang, "Tight bounds for some risk measures, with applications to robust portfolio selection," *Operations Research*, vol. 59, no. 4, pp. 847–865, 2011.

[24] E. Delage and Y. Ye, "Distributionally robust optimization under moment uncertainty with application to data-driven problems," *Operations Research*, vol. 58, no. 3, pp. 595–612, 2010.

[25] L. E. 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] M. J. Best and R. R. Grauer, "Sensitivity analysis for mean-variance portfolio problems," *Management Science*, vol. 37, no. 8, pp. 980–989, 1991.

[27] M. J. Best and R. R. Grauer, "On the sensitivity of mean-variance-efficient portfolios to changes in asset means: some analytical and computational results," *Review of Financial Studies*, vol. 4, no. 2, pp. 315–342, 1991.

[28] M. J. Best, *Portfolio Optimization*. CRC Press, 2010.

[29] A. Ben-Tal and A. Nemirovski, "Robust convex optimization," *Mathematics of Operations Research*, vol. 23, no. 4, pp. 769–805, 1998.

[30] A. Ben-Tal, T. Margalit, and A. Nemirovski, "Robust modeling of multi-stage portfolio problems," in *High Performance Optimization*, pp. 303–328, Springer, 2000.

[31] D. Goldfarb and G. Iyengar, "Robust portfolio selection problems," *Mathematics of Operations Research*, vol. 28, no. 1, pp. 1–38, 2003.

[32] L. Garlappi, R. Uppal, and T. Wang, "Portfolio selection with parameter and model uncertainty: A multi-prior approach," *Review of Financial Studies*, vol. 20, no. 1, pp. 41–81, 2007.

[33] A.-F. Ling and C.-X. Xu, "Robust portfolio selection involving options under a "marginal+ joint" ellipsoidal uncertainty set," *Journal of Computational and Applied Mathematics*, vol. 236, no. 14, pp. 3373–3393, 2012.

[34] K. Natarajan, M. Sim, and J. Uichanco, "Tractable robust expected utility and risk models for portfolio optimization," *Mathematical Finance*, vol. 20, no. 4, pp. 695–731, 2010.

[35] M. Ç. PıNar and R. H. Tütüncü, "Robust profit opportunities in risky financial portfolios," *Operations Research Letters*, vol. 33, no. 4, pp. 331–340, 2005.

[36] K. Schöttle and R. Werner, "Robustness properties of mean-variance portfolios," *Optimization*, vol. 58, no. 6, pp. 641–663, 2009.

[37] D. Bertsimas, D. B. Brown, and C. Caramanis, "Theory and applications of robust optimization," *SIAM Review*, vol. 53, no. 3, pp. 464–501, 2011.

[38] F. J. Fabozzi, P. N. Kolm, D. Pachamanova, and S. M. Focardi, *Robust Portfolio Optimization and Management*. John Wiley & Sons, 2007.

[39] F. J. Fabozzi, D. Huang, and G. Zhou, "Robust portfolios: contributions from operations research and finance," *Annals of Operations Research*, vol. 176, no. 1, pp. 191–220, 2010.

[40] D. A. Pachamanova, "Robust portfolio selection," *Wiley Encyclopedia of Operations Research and Management Science*, 2013.

[41] J. Mossin, "Optimal multiperiod portfolio policies," *The Journal of Business*, vol. 41, no. 2, pp. 215–229, 1968.

[42] D. Li and W.-L. Ng, "Optimal dynamic portfolio selection: Multiperiod mean-variance formulation," *Mathematical Finance*, vol. 10, no. 3, pp. 387–406, 2000.

[43] A. H. Chen, F. C. Jen, and S. Zionts, "The optimal portfolio revision policy," *Journal of Business*, pp. 51–61, 1971.

[44] E. J. Elton and M. J. Gruber, "On the optimality of some multiperiod portfolio selection criteria," *The Journal of Business*, vol. 47, no. 2, pp. 231–243, 1974.

[45] E. F. Fama, "Multiperiod consumption-investment decisions," *The American Economic Review*, vol. 60, no. 1, pp. 163–174, 1970.

[46] N. H. Hakansson, "Multi-period mean-variance analysis: Toward a general theory of portfolio choice," *The Journal of Finance*, vol. 26, no. 4, pp. 857–884, 1971.

[47] N. H. Hakansson, "On optimal myopic portfolio policies, with and without serial correlation of yields," *Journal of Business*, pp. 324–334, 1971.

[48] R. C. Merton, "Lifetime portfolio selection under uncertainty: The continuous-time case," *The Review of Economics and Statistics*, pp. 247–257, 1969.

[49] R. Östermark, "Vector forecasting and dynamic portfolio selection: Empirical efficiency of recursive multiperiod strategies," *European Journal of Operational Research*, vol. 55, no. 1, pp. 46–56, 1991.

[50] P. A. Samuelson, "Lifetime portfolio selection by dynamic stochastic programming," *The Review of Economics and Statistics*, pp. 239–246, 1969.

[51] R. L. Winkler and C. B. Barry, "A bayesian model for portfolio selection and revision," *The Journal of Finance*, vol. 30, no. 1, pp. 179–192, 1975.

[52] A. Takeda, S. Taguchi, and R. H. Tütüncü, "Adjustable robust optimization models for a nonlinear two-period system," *Journal of Optimization Theory and Applications*, vol. 136, no. 2, pp. 275–295, 2008.

[53] G. C. Calafiore, "Direct data-driven portfolio optimization with guaranteed shortfall probability," *Automatica*, vol. 49, no. 2, pp. 370–380, 2013.

[54] I. Popescu, "Robust mean-covariance solutions for stochastic optimization," *Operations Research*, vol. 55, no. 1, pp. 98–112, 2007.

[55] R. T. Rockafellar and S. Uryasev, "Optimization of conditional value-at-risk," *Journal of Risk*, vol. 2, pp. 21–42, 2000.

[56] R. T. Rockafellar and S. Uryasev, "Conditional value-at-risk for general loss distributions," *Journal of Banking & Finance*, vol. 26, no. 7, pp. 1443–1471, 2002.

[57] S. Zhu and M. Fukushima, "Worst-case conditional value-at-risk with application to robust portfolio management," *Operations Research*, vol. 57, no. 5, pp. 1155–1168, 2009.

[58] X. Tong, L. Qi, F. Wu, and H. Zhou, "A smoothing method for solving portfolio optimization with cvar and applications in allocation of generation asset," *Applied Mathematics and Computation*, vol. 216, no. 6, pp. 1723–1740, 2010.

[59] A. Ben-Tal and M. Teboulle, "An old-new concept of convex risk measures: The optimized certainty equivalent," *Mathematical Finance*, vol. 17, no. 3, pp. 449–476, 2007.

[60] F. Black and R. Litterman, "Global portfolio optimization," *Financial Analysts Journal*, vol. 48, no. 5, pp. 28–43, 1992.

[61] D. Goldfarb and G. Iyengar, "Robust portfolio selection problems," *Mathematics of Operations Research*, vol. 28, no. 1, pp. 1–38, 2003.

[62] S. Benartzi and R. H. Thaler, "Naive diversification strategies in defined contribution saving plans," *American economic review*, pp. 79–98, 2001.

[63] V. DeMiguel, L. Garlappi, and R. Uppal, "Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy?," *Review of Financial Studies*, vol. 22, no. 5, pp. 1915–1953, 2009.

[64] A. L. Gibbs and F. E. Su, "On choosing and bounding probability metrics," *International statistical review*, vol. 70, no. 3, pp. 419–435, 2002.

[65] G. C. Calafiore, "Ambiguous risk measures and optimal robust portfolios," *SIAM Journal on Optimization*, vol. 18, no. 3, pp. 853–877, 2007.

[66] E. Erdoğan and G. Iyengar, "Ambiguous chance constrained problems and robust optimization," *Mathematical Programming*, vol. 107, no. 1-2, pp. 37–61, 2006.

[67] M. Ç. Pınar and A. B. Paç, "Mean semi-deviation from a target and robust portfolio choice under distribution and mean return ambiguity," *Journal of Computational and Applied Mathematics*, vol. 259, pp. 394–405, 2014.

[68] H. Scarf, K. Arrow, and S. Karlin, "A min-max solution of an inventory problem," *Studies in the Mathematical Theory of Inventory and Production*, vol. 10, pp. 201–209, 1958.

[69] A. B. Paç and M. Ç. Pınar, "Robust portfolio choice with cvar and var under distribution and mean return ambiguity," *Top*, vol. 22, no. 3, pp. 875–891, 2014.

[70] A. Shapiro, "Topics in stochastic programming," *CORE Lecture Series, Universite Catholique de Louvain*, 2011.

[71] J. Owen and R. Rabinovitch, "On the class of elliptical distributions and their applications to the theory of portfolio choice," *The Journal of Finance*, vol. 38, no. 3, pp. 745–752, 1983.

[72] L. Chen, S. He, and S. Zhang, "When all risk-adjusted performance measures are the same: In praise of the sharpe ratio," *Quantitative Finance*, vol. 11, no. 10, pp. 1439–1447, 2011.