OPTIMIZATION WITH GENERALIZED DEVIATION MEASURES IN RISK MANAGEMENT

By

KONSTANTIN P. KALINCHENKO

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2012

I dedicate this thesis to my parents Pavel and Olga, and my brother Alexander, who supported me in all my endeavours.

ACKNOWLEDGMENTS

I am very thankful to my advisor Prof. Stan Uryasev and to Prof. R.Tyrrell Rockafellar for their support during my doctorate studying at the University of Florida. With their research guidance and invaluable help I was able to grow on both professional and personal levels. I would like to express my gratitude to other members of my doctorate committee, Prof. Panos Pardalos, Prof. Vladimir Boginski and Prof. Liqing Yan for their contribution to my research. Also, I am particularly thankful to Prof. Michael Zabarankin (Stevens Institute of Technology), Prof. Mark J. Flannery (University of Florida, Warrington College of Business Administration) and Prof. Oleg Bondarenko (University of Illinois) for valuable feedbacks. I would also like to express my greatest appreciation to my colleagues from the Risk Management and Financial Engineering lab and Center for Applied Optimization. Intensive discussions, exchange of ideas and joint research with my fellow graduate students from these two labs helped me significantly to achieve my goals. Also, I would like to thank my family and friends, who supported and encouraged me in all of my beginnings.

TABLE OF CONTENTS

LIST OF TABLES

Table page

LIST OF FIGURES

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

OPTIMIZATION WITH GENERALIZED DEVIATION MEASURES
IN RISK MANAGEMENT

By

Konstantin P. Kalinchenko

May 2012

Chair: Stan Uryasev
Major: Industrial and Systems Engineering

Our work provides an overview of the so-called generalized deviation measures and generalized risk measures, and develops stochastic optimization approaches utilizing them. These measures are designed to quantify risk when implied distributions are known. We provide useful examples of deviation and risk measures, which can be efficiently applied in situations, when the classical measures either do not properly account for risk, or do not satisfy properties desired for efficient application in stochastic optimization. We discuss the importance of considering alternative risk and deviation measures in the classical models, such as the capital asset pricing model and quantile regression. We apply stochastic optimization and risk management techniques based on the conditional value-at-risk (CVaR) to solve a dynamic sensor scheduling problem with robustness constraints on a wireless connectivity network. We also develop an efficient application of the generalized Capital Asset Pricing Model based on mixed CVaR deviation to estimating risk preferences of investors using S&P500 stock index option prices.

In the first part we provide an overview of the main classes of generalized deviation measures and corresponding risk measures, and compare them to the classical risk and deviation measures, such as maximum risk, value-at-risk and standard deviation. In addition, we provide a relation between deviation measures and measures of error, which are used in regression models. In some applications, such as simulation, a

distribution of the residual term has to be specified. We apply the entropy maximization principle to identify the appropriate distribution for the quantile regression (factor) model.

In the second part we consider several classes of problems that deal with optimizing the performance of dynamic sensor networks used for area surveillance, in particular, in the presence of uncertainty. The overall efficiency of a sensor network is addressed from the aspects of minimizing the overall information losses, as well as ensuring that all nodes in a network form a robust connectivity pattern at every time moment, which would enable the sensors to communicate and exchange information in uncertain and adverse environments. The considered problems are solved using mathematical programming techniques that incorporate CVaR, which allows one to minimize or bound the losses associated with potential risks. The issue of robust connectivity is addressed by imposing explicit restrictions on the shortest path length between all pairs of sensors and on the number of connections for each sensor (i.e., node degrees) in a network. Specific formulations of linear 0-1 optimization problems and the corresponding computational results are presented.

In the third part we apply the generalized Capital Asset Pricing Model based on mixed CVaR deviation to calibrate risk preferences of investors. We introduce the new generalized beta to capture tail performance of S&P500 returns. Calibration is done by extracting information about risk preferences from option prices on S&P500. Actual market option prices are matched with the estimated prices from the pricing equation based on the generalized beta. These results can be used for various purposes. In particular, the structure of the estimated deviation measure conveys information about the level of fear among investors. High level of fear reflects a tendency of market participants to hedge their investments and signals investors' anticipation of poor market trend. This information can be used in risk management and for optimal capital allocation.

CHAPTER 1
INTRODUCTION

The main objective of our study is to develop models utilizing generalized measures of deviation, risk and error in stochastic optimization and risk management applications. Uncertainty is generally modeled using random variables, and different models utilize various functionals (or measures) on the space of random variables to properly account for risk. Depending on a particular application, the functional must have certain properties to quantify a certain aspect of uncertainty. Although many different functionals may satisfy these properties, most models utilize just several classic functionals, such as standard deviation or quantile.

Optimality of any particular choice of measure accounting for uncertainty can often be argued. This led to multiple studies introducing new measures and developing alternative models utilizing them. The concepts of generalized measures of deviation, risk and error were developed to wrap these and other measures in several classes, where each class satisfies certain properties (axioms) required in a particular application. If a certain model based on some risk or deviation functional is adjusted to certain assumptions, its application can often be generalized by substituting the functional with a generalized measure of risk or deviation. Different instances of the generalized model can thus be compared. Moreover, if the functional has a parameter, it can also be optimized.

In our work, we utilize three classes of functionals: generalized measures of risk, generalized deviation measures, and generalized measures of error. Generalized measures of risk were designed to quantify potential losses. Generalized deviation measures account only for variability of losses. Generalized measures of error can be viewed as tools to estimate significance of a residual term in approximation, or its deviation from 0.

It is important to mention that there is a one-to-one correspondence between measures of risk and measures of deviation, and every measure of error has a particular measure of deviation corresponding to it. This feature links together certain very different models, and provides a solution to properly choosing functional in one model depending on the functional used in another model, when the two models are applied to the same problem.

Although the theoretical foundations of the deviation measures have already been developed, their practical applications have not yet become popular. In our work, we demonstrate several applications of the generalized measures of deviation, risk and error in stochastic optimization. In particular, we consider instances based on conditional value-at-risk (CVaR) and mixed CVaR. CVaR is a risk measure, which gained substantial attention in academic publications due to several reasons. First, CVaR has an intuitive definition as expected losses corresponding to the $1 - \alpha$ tail of distribution. Second, CVaR is a coherent measure of risk, and is therefore applicable in optimization. Third, the problem of optimizing CVaR has a linear programming formulation. Mixed CVaR is a convex combination of several CVaR terms with different α values. By varying the number of terms and the values of coefficients and α, one can precisely specify significance of different parts of a distribution according to his perception of risk.

In Chapter 2 we applied entropy maximization methodology to specifying distribution of the residual term in generalized linear regression. Generalized linear regression is defined as a stochastic optimization problem of minimizing a generalized measure of error of the residual. It is important to mention that the generalized linear regression has an alternative formulation utilizing deviation measure and so-called statistic, both corresponding to the same measure of error. Two instances of the generalized linear regression are well known: classical linear regression, based on mean squared error, and quantile regression, based on Koenker-Bassett error. In certain applications, such as simulation, the distribution of the residual term in the linear regression has to be specified. A common way to specify the distribution under limited information is by maximizing Shannon entropy. This approach is justified by the common view that entropy is a measure of uncertainty. In the case of the classical linear regression, when expectation and variance of the residual term are known, the distribution with maximum entropy is normal. Following the same intuition, in quantile regression we estimated the distribution by maximizing entropy subject to constraints on quantile and CVaR deviation.

In Chapter 3 we applied CVaR-based optimization to a sensor scheduling problem. Such problems are common in applications where information losses occur due to inability to collect information from all sources simultaneously. Information losses associated with not observing a certain site at some moment in time are modeled as a penalty, which consists of two components: a fixed penalty and a penalty proportional to the time the site was not observed. In this setup CVaR is applied to minimize the average of the $1-\alpha$ greatest penalties. The model also includes two types of wireless connectivity robustness constraints: 2-club and k-plex.

We discuss an example of the problem requiring optimization of a deviation measure in Chapter 4. We considered the generalized Capital Asset Pricing Model based on mixed CVaR deviation to estimate risk preferences of investors. The problem of risk preferences estimation was actively discussed in many studies. One motivation for these discussions is the criticism of the classical CAPM, which is based on the assumption that investors' perception of risk can be represented by standard deviation. This criticism aligns with our motivation for applying the concept of generalized measures. Contrary to the classical CAPM and some recent modifications, generalized CAPM considers a class of mixed CVaR deviations with coefficients specifying parameterization.

CHAPTER 2
GENERALIZED MEASURES OF DEVIATION, RISK AND ERROR

In this chapter we provide an overview of generalized deviation measures[1] and related quantitative measures, e.g. risk measures, measures of error, statistics and entropy. Most of these measures have been introduced in a recent line of research by R.T.Rockafellar, S.Uryasev, M.Zabarankin and S.Sarykalin. We demonstrate that some subclasses of these measures have properties, which allow them to be more efficiently applied in risk management applications.

An application of the deviation measures in regression models is proposed in this chapter. In particular, we identified the distribution which is the most applicable to model the residual in the quantile regression.

The following section provides an overview of the popular measures used in risk management. Section 3 introduces so-called generalized deviation measures and generalized risk measures, and provides mathematical relations between them. Some important subclasses of these measures are also introduced in Section 3. An overview of conditional value-at-risk deviation and related measures is provided in section 4. Section 5 contains an overview of so-called measures of error and their relation to deviation measures. The same section introduces applications of the deviation measures and the measures of error in regression models.

2.1 Classical Risk and Deviation Measures

Since risk management became a standard practice for almost all institutions and commercial companies, a number of quantitative measures have been developed to

[1] We use the word ”generalized” to accentuate that these measures belong to a class generalizing the standard deviation. In the context of this and the following chapters the terms ”deviation measure” and ”generalized deviation measure” have the same meaning.

provide a natural way to estimate risk. The most popular and commonly used measures have been maximum risk, value-at-risk and standard deviation.

Maximum risk (*maxrisk*) provides a quantitative evaluation of losses in the worst possible scenario. The expression below provides a formal definition of *maxrisk*:

$$\mathrm{maxrisk}(X) = -\inf(X)$$

It is the least convenient measure due to its over-conservatism: for most common distributions it provides a meaningless value of infinity. Even when this measure is applicable (e.g. when a set of possible outcomes is finite), this measure lacks robustness. For example, if the set of possible outcomes is extended by adding one more outcome corresponding to losses, which are substantially greater than the previous value of *maxrisk*, then the value of *maxrisk* changes by the same amount regardless of the probability of this outcome. It is important to mention, however, that this measure can be efficiently used in some optimization applications. In particular, any constraint on *maxrisk* is equivalent to a set of similar constraints for each possible outcome. This was demonstrated, for example, in Boyko et al. (2011).

Value-at-risk has been a popular measure in the last 20 years due to its natural interpretation as an amount of reserves required to prevent default with a given probability. Below is the formal definition from Artzner et al. (1999):

Definition. Given $\alpha \in (0, 1)$, and a reference instrument r, the value-at-risk VaR_α at level α of the final net worth X with distribution P, is defined by the following relation:

$$\mathrm{VaR}_\alpha(X) = -\inf\{\, x \mid P[X \leq x \cdot r] > \alpha \,\}$$

Basel Committee on Banking Supervision (BCBS) issues so-called Basel Accords, which are recommendations on banking laws and regulations. According to these recommendations, value-at-risk is the preferred approach to market risk measurement (for example, Bas (2004)). In particular, these recommendations specify minimum capital requirements, estimated with value-at-risk. In many countries (including the USA) regulators enforce financial companies to comply with some or all of these recommendations. This led to value-at-risk becoming one of the most commonly used risk measures.

Despite its popularity, value-at-risk has a serious drawback. The problem is that the functional VaR_α does not have a convexity property. In risk management convexity is often a necessary requirement. For example, in portfolio management, convexity of a risk measure justifies diversification of investments. Also, the lack of convexity makes the value-at-risk measure inefficient in optimization, where convexity of the optimized function or constraints is always a desired property.

Standard deviation $\sigma(\cdot)$ is a function defined on the space of random variables as a square root of variance. In some applications, a term volatility is used instead of standard deviation. Contrary to value-at-risk, standard deviation satisfies the convexity property. This measure is very useful, because, in particular, volatility in many models is a parameter. For example, in portfolio management a stock price random process S_t is described by the following stochastic differential equation:

$$dS_t = \mu_t S_t dt + \sigma_t S_t dW_t$$

with σ_t denoting stock volatility.

Volatility and variance can be viewed as the most popular measures in portfolio optimization and risk management. For example, the Capital Asset Pricing Model (CAPM, Sharpe (1964), Lintner (1965), Mossin (1966), Treynor (1961), Treynor (1999)) and the Arbitrage Pricing Theory (APT, Ross (1976)) are factor models focusing on explaining variability in stock returns. CAPM assumptions imply, in particular, that all investors in the market are optimizing their investment portfolios considering variance as a measure of risk. Based on variance, CAPM introduces the so-called Beta, a quantity determining exposure of stock (or portfolio) returns to future market trend:

$$\beta_i = \frac{\sigma_{iM}}{\sigma_M^2}$$

where σ_{iM} denotes covariance between stock i returns and market returns, and σ_M denotes the market volatility. If all CAPM assumptions hold, the total variance of a stock return can be separated into systematic and nonsystematic (idiosyncratic) components, where systematic part of the variance corresponds to market returns:

$$\sigma_i^2 = \beta_i^2 \sigma_M^2 + \sigma_{i,n}^2$$

In the expression above β_i denotes the Beta of the stock i, and $\sigma_{i,n}$ is the nonsystematic (stock-specific) part of volatility. This expression provides a tool for risk management. For example, if a manager is looking for a portfolio with weights (w_i) with no exposure to market trend, he has to consider only portfolios with total Beta equal 0:

$$\sum_{i=1}^{N} w_i \beta_i = 0$$

The convexity property of volatility guarantees that the nonsystematic component of the portfolio variance can be reduced by diversification.

It is also important to mention that for a normal random variable X, a pair $(E[X], \sigma(X))$ provides complete information about the distribution of X. Moreover, assume that the distribution of Y has to be specified, and the only available information about a random variable Y is the values $\mu_y = E[Y]$ and $\sigma_y = \sigma(Y)$. Following the maximum entropy principle, which was first introduced in Jaynes (1957, 1968), it is natural to assume the least-informative distribution of Y with given mean and variance. Specifically, consider the following optimization problem:

$$\begin{aligned} \max \quad & \text{Entr}(f) \\ \text{s.t.} \quad & \int_{-\infty}^{\infty} t f(t) dt = \mu \end{aligned}$$

$$\int_{-\infty}^{\infty} t^2 f(t)dt - \mu^2 = \sigma^2$$
$$f \text{ is a PDF}$$

where $\mathrm{Entr}(f)$ denotes the Shannon entropy (Shannon (1948)):

$$\mathrm{Entr}(f) = -\int_{-\infty}^{\infty} f(t) \ln f(t) dt$$

which is a common measure of uncertainty. The optimal solution $f^*(t) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(t-\mu)^2}{2\sigma^2}}$ is the probability density function of a normal distribution with mean μ and variance σ^2 (the proof can be found in Cozzolino and Zahner (1973)). Therefore, if the only available information about a distribution is mean and variance, in can be natural to assume a normal distribution. It is very convenient, because in many models, such as factor models, uncertainty is modeled by normal distribution.

The standard deviation, however, has disadvantages. In particular, it doesn't satisfy the monotonicity property

$$X < Y \ a.s. \implies R(X) < R(Y)$$

It would be natural to assume that a risk measure should be monotonic. For example, if the probability space for a given random variable consists of only one event, even if it is associated with big losses, the value of the standard deviation equals the lowest possible value 0. Therefore, in the framework of our work, we assume that the standard deviation measure is not a risk measure, but instead belongs to a class of (generalized) deviation measures, which will be defined in the next section.

Another drawback of the standard deviation is that it doesn't distinguish between desirable and undesirable parts of the distribution, and it is not sensitive to rare but critical outcomes, corresponding to one tail of the distribution. Empirical studies demonstrate that, for example, distributions of stock returns have heavy (negative) tails associated with extreme market conditions (an overview of heavy tails in financial risk can be found in Bradley and Taqqu (2003)). The worst-case $p\%$ outcomes particularly interest investors, because they are the ones most likely to cause a default.

The inefficiencies of the standard measures, mentioned above, can be summarized as follows:

1. The classical value-at-risk measure lacks the convexity property, therefore it is sometimes difficult to implement VaR in stochastic optimization.
2. The maximum risk measure can be easily implemented in stochastic optimization, but its over-conservatism often makes it meaningless.
3. The standard deviation and the variance do not satisfy the monotonicity property, which would be natural for a risk measure.
4. As a deviation measure, the standard deviation does not distinguish between the positive outcomes (gains) and negative outcomes (losses), and measures overall variability, while risk manager is primarily concerned about the part of variability associated with the most undesirable scenarios.

2.2 Generalized Risk and Deviation Measures

A new systematization of measures evaluating probability distributions was introduced by Rockafellar, Sarykalin, Uryasev and Zabarankin (Rockafellar et al. (2006a), Sarykalin (2008)). They introduce a number of axioms defining two separate classes: deviation measures and risk measures. They also provide a one-to-one correspondence between these classes and special subclasses.

Consider the following set of axioms:

(D1) $\mathcal{D}(X+C) = \mathcal{D}(X)$ for all X and constants C,

(D2) $\mathcal{D}(0) = 0$ and $\mathcal{D}(\lambda X) = \lambda\mathcal{D}(X)$ for all X and all $\lambda > 0$,

(D3) $\mathcal{D}(X+Y) \leq \mathcal{D}(X) + \mathcal{D}(Y)$ for all X and Y,

(D4) $\mathcal{D}(X) \geq 0$ for all X with $\mathcal{D} > 0$ for nonconstant X,

(D5) $\{\, X \mid \mathcal{D}(X) \leq C \,\}$ is closed for every constant C,

(D6) $\mathcal{D}(X) \leq EX - \inf X$ for all X.

The axiom (D2) defines positive homogeneity, axioms (D2) and (D3) together define convexity. According to the definition in Rockafellar et al. (2006a), a functional

$\mathcal{D} : \mathcal{L}^2 \to [0, \infty]$ is a deviation measure if it satisfies axioms (D1)-(D4). The axiom (D5) defines closedness. In this document we will only consider closed deviation measures. The property (D6) defines lower range dominance.

Under these axioms, $\mathcal{D}(X)$ depends only on $X - EX$ (from the case of (D1) where $C = -EX$), and it vanishes only if $X - EX = 0$ (as seen from (D4) with $X - EX$ in place of X. This captures the idea that $\mathcal{D}$ measures the degree of uncertainty in X. Proposition 4 in Rockafellar et al. (2006a) proves the convexity property of the class of deviation measures and the subclass of lower range dominated deviation measures. It can be seen that the standard deviation fits into the class of deviation measures, but it doesn't satisfy the lower range dominance axiom (D6).

Although the deviation measures as measures of uncertainty provide some information about the riskiness associated with the outcome of X, they are not risk measures in the sense proposed in Artzner et al. (1999). Consider, for example, a situation in the financial market with an arbitrage opportunity with a net payoff X. By definition of the arbitrage, $X \geq 0$ almost surely and $P(X > 0) > 0$. Arbitrage is generally viewed as a profitable riskless opportunity, therefore, for a risk measure $\mathcal{R}$ the value $\mathcal{R}(X)$ should not be greater than 0. If X is random, a deviation measure will always be greater than 0.

Rockafellar et al. (2006a) introduces the class of coherent risk measures, which extends the risk measures defined in Artzner et al. (1999). Consider the following axioms:

(R1) $\mathcal{R}(X + C) = \mathcal{R}(X) - C$ for all X and constants C,

(R2) $\mathcal{R}(0) = 0$, and $\mathcal{R}(\lambda X) = \lambda \mathcal{R}(X)$ for all X and constants $\lambda > 0$,

(R3) $\mathcal{R}(X + Y) \leq \mathcal{R}(X) + \mathcal{R}(Y)$ for all X and Y,

(R4) $\mathcal{R}(X) \leq \mathcal{R}(Y)$ for all $X \geq Y$,

(R5) $\mathcal{R}(X) > E[-X]$ for all non-constant X.

The axiom (R2) defines positive homogeneity, (R3) defines subadditivity. The axioms (R2) and (R3) combined imply that $\mathcal{R}$ is a convex functional. (R4) is called monotonicity.

According to the definitions in Rockafellar et al. (2006a), a functional $\mathcal{R} : \mathcal{L}^2 \to (-\infty, \infty]$ is

a) a coherent risk measure if it satisfies axioms (R1)-(R4),

b) strictly expectation bounded risk measure if it satisfies (R1)-(R3) and (R5), and

c) coherent, strictly expectation bounded risk measure if it satisfies all axioms (R1)-(R5).

It can be noticed that VaR doesn't fit in any of these categories because it doesn't have the subadditivity property (R3).

The same paper defines relations between deviation measures and risk measures:

$$\mathcal{D}(X) = \mathcal{R}(X - EX) \tag{2–1}$$

$$\mathcal{R}(X) = E[-X] + \mathcal{D}(X) \tag{2–2}$$

In particular, equations (2–1) and (2–2) provide one-to-one correspondence between strictly expectation bounded risk measures (satisfying (R1)-(R3) and (R5)) and deviation measures (satisfying axioms (D1)-(D4)), and one-to-one correspondence between coherent, strictly expectation bounded risk measures (satisfying (R1)-(R5)) and lower range dominated deviation measures (satisfying (D1)-(D4) and (D6)). It can be shown directly or using these relations that the sets of coherent risk measures, strictly expectation bounded risk measures, and coherent, strictly expectation bounded risk measures are all convex.

According to these relations, the standard deviation σ corresponds to the strictly expectation bounded risk measure $\mathcal{R}_\sigma(x) = E[-X] + \sigma(X)$, which is not a coherent risk measure.

2.3 Conditional Value-at-Risk

The classical risk measures discussed in the beginning of the chapter have a number of imperfections. This led to the development of new kinds of risk measures. One of the most noticeable risk measures is the conditional value-at-risk (CVaR), also known as expected shortfall, or tail-VaR. We define this measure according to Pflug (2000):

$$\mathrm{CVaR}_\alpha(X) = \min_C \left\{ C + (1-\alpha)^{-1} E[X + C]_- \right\}$$

where $[y]_-$ equals $-y$ for $y < 0$ and 0 otherwise. The parameter α must have a value in the interval $(0, 1)$. It is important to notice that the optimal C equals value-at-risk:

$$\mathrm{argmin}_C \left\{ C + (1-\alpha)^{-1} E[X + C]_- \right\} = \mathrm{VaR}_\alpha(X)$$

where VaR denotes the value-at-risk:

$$\mathrm{VaR}_\alpha(X) = -\sup\{\, z \mid F_X(z) < 1 - \alpha \,\}$$

An equivalent definition can be found in Acerbi (2002):

$$\mathrm{CVaR}_\alpha(X) = (1-\alpha)^{-1} \int_\alpha^1 \mathrm{VaR}_\beta(X)\, d\beta$$

If CVaR_α is continuous at $-\mathrm{VaR}_\alpha$, then it can be expressed via the following relation:

$$\mathrm{CVaR}_\alpha(X) = -E[X \mid X \le -\mathrm{VaR}_\alpha(X)] \tag{2–3}$$

Equation (2–3) shows that conditional value-at-risk equals the weighted average of losses exceeding value-at-risk. Therefore, conditional value-at-risk estimates how severe can be potential losses associated with a tail of distribution. This can be viewed as an advantage over the classical value-at-risk, which is not sensitive to changes in the tail of distribution. Also, conditional value-at-risk is a coherent, strictly expectation bounded risk measure. This feature allows considering CVaR in a variety of stochastic optimization applications. In particular, as it will be discussed in Chapter 3, conditional value-at-risk can be used in linear programming. The lower range dominated deviation measure, defined via (2–1) for $\mathcal{R} = \text{CVaR}_\alpha$ is called a conditional value-at-risk deviation (CVaR deviation), denoted as $\text{CVaR}^\Delta_\alpha$. Due to convexity of coherent, strictly expectation bounded risk measures, a convex combination of several CVaRs with different confidence levels α is also a coherent strictly expectation bounded risk measure. Convex combination of several CVaRs is called a mixed conditional value-at-risk (mixed CVaR). The lower range dominated deviation measure, defined via (2–1) for $\mathcal{R}$ = *mixed CVaR*, is called a mixed conditional value-at-risk deviation (mixed CVaR deviation). In Chapter 4 we demonstrate application of mixed CVaR deviation in the framework of the generalized capital asset pricing model for estimating risk preferences of investors.

2.4 Application to Generalized Linear Regressions

In this section we provide an overview of the so-called measures of error, including their relation to generalized deviation measures and application to generalized linear regressions.

2.4.1 Measures of Error

Consider a functional $\mathcal{E} : \mathcal{L}^2 \to [0, \infty]$ and the set of axioms:

(E1) $\mathcal{E}(0) = 0$ but $\mathcal{E}(X) > 0$ when $X \neq 0$; also, $\mathcal{E}(C) < \infty$ for constants C,

(E2) $\mathcal{E}(\lambda X) = \lambda \mathcal{E}(X)$ for constants $\lambda > 0$,

(E3) $\mathcal{E}(X + Y) \leq \mathcal{E}(X) + \mathcal{E}(Y)$ for all X, Y,

(E4) $\{ X \in \mathcal{L}^2(\Omega) \,|\, \mathcal{E} \leq c \}$ is closed for all $C < \infty$,

(E5) $\inf_{X \,:\, EX = C} \mathcal{E}(X) > 0$ for constants $C \neq 0$.

According to the definition in Rockafellar et al. (2008), the functional $\mathcal{E}$ is a measure of error if it satisfies axioms (E1)-(E4). The property (E5) is called nondegeneracy.

Consider a functional $\mathcal{D} : \mathcal{L}^2 \to [0, \infty]$ defined according to the following relation:

$$\mathcal{D}(X) = \inf_C \mathcal{E}(X - C) \tag{2–4}$$

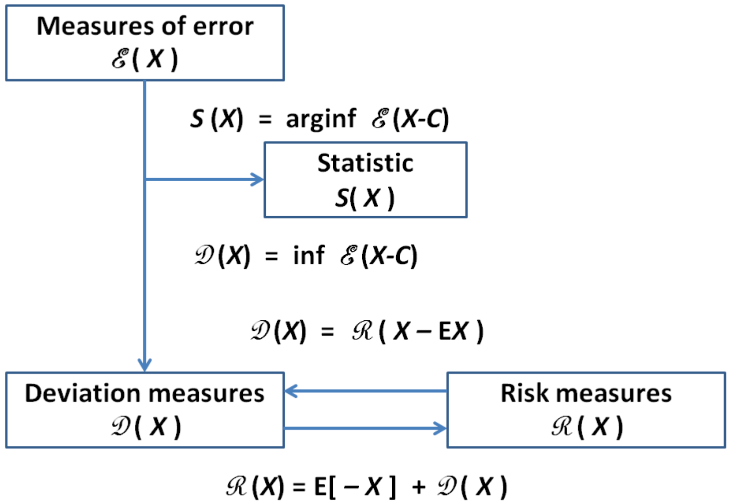

Figure 2-1. Relations Between Measures of Error, Deviation Measures, Risk Measures and Statistics

where $\mathcal{E}$ is a nondegenerate measure of error. Then, according to Theorem 2.1 in Rockafellar et al. (2008), $\mathcal{D}$ is a deviation measure. We will further use the term projected deviation measure to specify that this deviation measure is obtained according to equation (2–4). The set $S(X) = \text{argmin}_C \mathcal{E}(X - C)$ is called statistic. In many cases $S(X)$ is reduced to a single value.

Consider a mean square error: $\mathcal{E}_{MS}(X) = E[X^2]$. It is a well-known fact that $S_{MS}(X) = EX$. According to (2–4), the corresponding deviation measure is variance: $\mathcal{D}_{MS} = \sigma^2$.

Figure 2-1 illustrates relations between measures of error, statistics, deviation measures and risk measures.

Another important example is the Koenker-Bassett error, introduced in Koenker and Bassett (1978):

$$\mathcal{E}_{KB} = E\left[X_+ + \frac{\alpha}{1-\alpha}X_-\right]$$

where $X_+ = \max\{0, X\}$ and $X_- = \max\{0, -X\}$. It was shown in Rockafellar et al. (2008) that $\mathcal{E}_{KB}$ is a nondegenerate measure of error, which corresponds to the conditional value-at-risk deviation:

$$\mathcal{D}_{KB}(X) = \inf_{C} \mathcal{E}_{KB}(X - C) = \text{CVaR}_{\alpha}^{\Delta}(X)$$

Moreover, the corresponding $S_{KB}(X)$ equals value-at-risk:

$$S_{KB}(X) = \text{argmin}_{C} \mathcal{E}_{KB}(X - C) = \text{VaR}_{\alpha}(X)$$

2.4.2 Generalized Linear Regressions

Generalized linear regression was defined in Rockafellar et al. (2008) as the following problem:

$$\min \quad \mathcal{E}(Z(c_0, c_1, \dots, c_n)) \tag{2–5}$$

$$\text{s.t.} \quad Z(c_0, c_1, \dots, c_n) = Y - (c_0 + c_1 X_1 + \dots + c_n X_n) \tag{2–6}$$

In the above formulation the random variables $X_1, \dots, X_n$ are factors and $c_0, c_1, \dots, c_n$ are regression coefficients. If $\mathcal{E}(Y) < \infty$, then the set of optimal regression coefficients $(\bar{c}_0, \bar{c}_1, \dots, \bar{c}_n)$ always exists.

Theorem 3.2 in the same paper proves equivalence of the problem (2–5) — (2–6) to the following problem of minimizing the projected deviation measure:

$$\min \quad \mathcal{D}(Z(c_0, c_1, \dots, c_n)) \tag{2–7}$$

$$\text{s.t.} \quad 0 \in S(Z(c_0, c_1, \dots, c_n)) \tag{2–8}$$

$$Z(c_0, c_1, \dots, c_n) = Y - (c_0 + c_1 X_1 + \dots + c_n X_n) \tag{2–9}$$

For the optimal regression coefficients $\bar{c}_0, \bar{c}_1, \dots, \bar{c}_n$ we can write:

$$Y = \bar{c}_0 + \bar{c}_1 X_1 + \dots + \bar{c}_n X_n + \varepsilon \tag{2–10}$$

where ε is the error term, equal to the residual for the optimal regression coefficients:

$$\varepsilon = Z(\bar{c}_0, \bar{c}_1, \dots, \bar{c}_n)$$

The Theorem 3.2 and equation (2–10) imply that for the optimal set of coefficients statistic of Y equals statistic of the optimal combination of factors:

$$S(Y) = S(\bar{c}_0 + \bar{c}_1 X_1 + \ldots + \bar{c}_n X_n)$$

They also imply that the optimal set of coefficients minimizes the deviation measure of the residual $Z(c_0, c_1, \ldots, c_n)$:

$$\mathcal{D}(\varepsilon) = \min_{c_0, c_1, \ldots, c_n} \mathcal{D}(Z(c_0, c_1, \ldots, c_n)) \tag{2–11}$$

It is important to notice that the value of c_0 has no effect on the right hand side of equation (2–11).

Below we consider two important examples, which will be used in further analysis.

Consider the mean squared error $\mathcal{E}_{MS}$. For this measure of error the problem (2–5) — (2–6) is equivalent to the classical linear regression. $S_{MS}(\cdot) = E[\cdot]$ implies that the expectation of the optimal combination of factors equals expectation of Y:

$$EY = E[\bar{c}_0 + \bar{c}_1 X_1 + \ldots + \bar{c}_n X_n]$$

Equation (2–11) implies that the error ε can be interpreted as a random variable, minimizing residual variance:

$$\sigma^2(\varepsilon) = \min_{c_1, \ldots, c_n} \sigma^2(Y - c_1 X_1 - \ldots - c_n X_n)$$

Another example is the quantile regression (for example, Koenker (2005)). It can be formulated according to (2–5) — (2–6) with $\mathcal{E} = \mathcal{E}_{KB}$, or, equivalently, according to (2–7) — (2–9) with a generalized deviation measure $\mathcal{D} = \text{CVaR}_\alpha^\Delta$ and statistic $S = \text{VaR}_\alpha$. Following the same logic that we used for the classical linear regression, $S_{KB}(\cdot) = \text{VaR}_\alpha(\cdot)$ implies the following:

$$\text{VaR}_\alpha(Y) = \text{VaR}_\alpha(\bar{c}_0 + \bar{c}_1 X_1 + \ldots + \bar{c}_n X_n)$$

Equation (2–11) for quantile regression implies the following interpretation of the optimal residual ε:

$$\text{CVaR}_\alpha^\Delta(\varepsilon) = \min_{c_1,\dots,c_n} \text{CVaR}_\alpha^\Delta(Y - c_1X_1 - \dots - c_nX_n)$$

2.4.3 Distribution of Residual

In some applications it is convenient to specify the distribution of the residual error ε. The choice of the distribution should be only based on the available information. For the generalized linear regression, only the statistic and deviation measure of the error term are known. Given this information, it is natural to pick the distribution which has the greatest uncertainty.

We consider the Shannon entropy $\text{Entr}(f)$, which is commonly used as a measure of uncertainty (Shannon (1948)):

$$\text{Entr}(f) = -\int_{-\infty}^{\infty} f(t)\ln f(t)dt$$

where f is the probability density function.

For convenience, consider the following notation. For a functional $\mathcal{F} : \mathcal{L}^2 \to [-\infty, \infty]$, define

$$\mathcal{F} : \{ f \,|\, f \text{ is a PDF } \} \to [-\infty, \infty]$$

according to the following:

$$\mathcal{F}(f) = \mathcal{F}(X) \text{ for } X \text{ such that } f \text{ is a PDF of } X$$

We apply this notation for $\mathcal{F}(\cdot) = E[\cdot]$, $\mathcal{F}(\cdot) = \sigma(\cdot)$, $\mathcal{F}(\cdot) = \text{CVaR}_\alpha^\Delta(\cdot)$, $\mathcal{F}(\cdot) = \text{VaR}_\alpha(\cdot)$.

For a generalized linear regression, defined in (2–7) — (2–8), we choose the distribution f_ε by solving the following entropy maximization problem:

$$\max \quad \text{Entr}(f) \tag{2–12}$$

$$\text{s.t.} \quad S(f) = 0 \tag{2–13}$$

$$\mathcal{D}(f) = \mathcal{D}(\varepsilon) \tag{2–14}$$

$$f \text{ is a PDF} \tag{2–15}$$

where $S(f)$ and $\mathcal{D}(f)$ are the statistic and the deviation measure of a random variable with the probability density function f.

For the classical linear regression the optimization problem (2–12) — (2–15) is the entropy maximization problem with constraints on expectation and expected value. The solution to this problem is the normal distribution.

For the quantile regression the optimization problem (2–12) — (2–15) is equivalent to the following:

$$\max \quad \text{Entr}(f) \tag{2–16}$$

$$\text{s.t.} \quad \text{VaR}_\alpha(f) = 0 \tag{2–17}$$

$$\text{CVaR}^\Delta_\alpha(f) = \text{CVaR}^\Delta_\alpha(\varepsilon) \tag{2–18}$$

$$f \text{ is a PDF} \tag{2–19}$$

We derive the solution to this problem from the solution to a similar problem:

$$\max \quad \text{Entr}(g) \tag{2–20}$$

$$\text{s.t.} \quad E(g) = 0 \tag{2–21}$$

$$\text{CVaR}^\Delta_\alpha(g) = v \tag{2–22}$$

$$g \text{ is a PDF} \tag{2–23}$$

The solution to (2–20) — (2–23) for $v = 1$ can be found in Grechuk et al. (2009):

$$g_{\varepsilon,1}(t) = \begin{cases} (1-\alpha)\exp\left(\frac{1-\alpha}{\alpha}\left(t - \frac{2\alpha-1}{1-\alpha}\right)\right) & t \le \frac{2\alpha-1}{1-\alpha} \\ (1-\alpha)\exp\left(-\left(t - \frac{2\alpha-1}{1-\alpha}\right)\right) & t \ge \frac{2\alpha-1}{1-\alpha} \end{cases}$$

Derivation of the solution for other values of v is identical to the case $v = 1$. The following function is the optimal probability density function:

$$g_{\varepsilon,v}(t) = \begin{cases} \frac{1-\alpha}{v} \exp\left(\frac{1-\alpha}{v\alpha}\left(t - \frac{2\alpha-1}{1-\alpha}\right)\right) & t \le \frac{2\alpha-1}{1-\alpha} \\ \frac{1-\alpha}{v} \exp\left(-\frac{1}{v}\left(t - \frac{2\alpha-1}{1-\alpha}\right)\right) & t \ge \frac{2\alpha-1}{1-\alpha} \end{cases} \tag{2–24}$$

Note, that

$$\mathrm{VaR}_{\alpha}(g_{\varepsilon,v}) = -\frac{2\alpha-1}{1-\alpha} \tag{2–25}$$

what follows from (2–24) and the following equality:

$$\int_{-\infty}^{\frac{2\alpha-1}{1-\alpha}} \frac{1-\alpha}{v} \exp\left(\frac{1-\alpha}{v\alpha}\left(t - \frac{2\alpha-1}{1-\alpha}\right)\right) dt = 1-\alpha$$

Theorem. The optimal probability density function f in the problem (2–16) — (2–19) is the following function $f_{\varepsilon,v}$:

$$f_{\varepsilon,v}(t) = \begin{cases} \frac{1-\alpha}{v} \exp\left(\frac{1-\alpha}{v\alpha} t\right) & t \le 0 \\ \frac{1-\alpha}{v} \exp\left(-\frac{1}{v} t\right) & t \ge 0 \end{cases} \tag{2–26}$$

where $v = \mathrm{CVaR}_{\alpha}^{\Delta}(\varepsilon)$.

Proof of Theorem. First, consider the following entropy property:

$$\mathrm{Entr}(f) = \mathrm{Entr}(g) \text{ for } g(t) = f(t-c)$$

where c is any constant. This property guarantees that the problem (2–16) — (2–19) is equivalent to the problem:

$$\max \quad \mathrm{Entr}(g) \tag{2–27}$$

$$\text{s.t.} \quad g(t) = f(t + E(f)) \ \forall t \tag{2–28}$$

$$\mathrm{VaR}_{\alpha}(f) = 0 \tag{2–29}$$

$$\mathrm{CVaR}_{\alpha}^{\Delta}(f) = v \tag{2–30}$$

$$f, g \text{ are PDFs} \tag{2–31}$$

The axiom (D1) allows us to substitute constraint (2–30) with constraint $\text{CVaR}_\alpha^\Delta(g) = d$. Also note: (2–28) guarantees that $E(g) = 0$ for the optimal g, and (2–29) together with (2–28) guarantee that $f(t) = g(t - \text{VaR}_\alpha(g))$. Therefore, this problem is also equivalent to the following:

$$\max \quad \text{Entr}(g) \tag{2–32}$$

$$\text{s.t.} \quad g(t) = f(t + E(f)) \ \forall t \tag{2–33}$$

$$f(t) = g(t - \text{VaR}_\alpha(g)) \ \forall t \tag{2–34}$$

$$\text{VaR}_\alpha(f) = 0 \tag{2–35}$$

$$\text{CVaR}_\alpha^\Delta(g) = v \tag{2–36}$$

$$E(g) = 0 \tag{2–37}$$

$$f, g \text{ are PDFs} \tag{2–38}$$

Note that in this problem the constraints (2–33), (2–34) and (2–35) are redundant: for any g in the feasible set, if f is defined according to (2–34), then such f already satisfies (2–35), and (2–33) is enforced by (2–37). Therefore, this problem is equivalent to the problem (2–20) — (2–23). $g_{\varepsilon,d}$ in (2–24) is the optimal PDF g, and the optimal f is obtained directly from (2–25) and (2–34). $\diamond$

This distribution has parameters α and v. We will further say that a random variable ε is distributed according to $\text{DExp}(\alpha, v)$, if its probability density function f_ε is expressed by (2–26).

Figures 2-2 and 2-3 depict the probability density functions of distributions $\text{DExp}(\alpha, v)$ for different values α and v. Each distribution can be viewed as a "two-sided exponential" distribution.

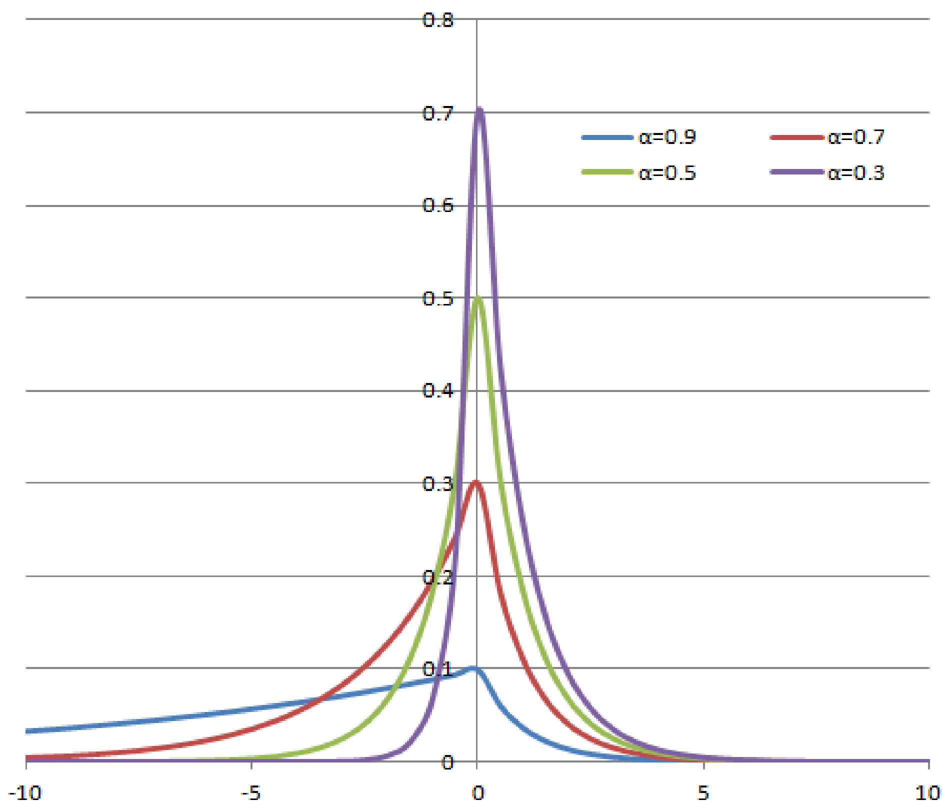

Figure 2-2. Probability Density Functions for $\mathrm{DExp}(\alpha, 1)$

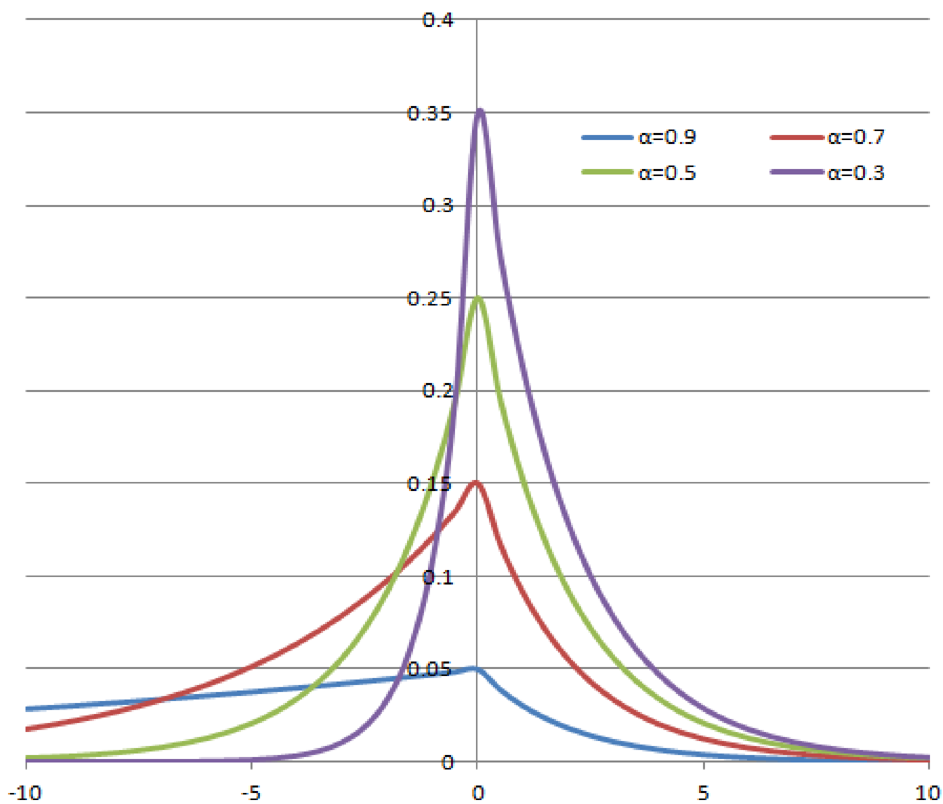

Figure 2-3. Probability Density Functions for $\mathrm{DExp}(\alpha, 2)$

CHAPTER 3
ROBUST CONNECTIVITY ISSUES IN DYNAMIC SENSOR NETWORKS FOR AREA SURVEILLANCE UNDER UNCERTAINTY

In this chapter[1] , we address several problems and challenges arising in the task of utilizing dynamic sensor networks for area surveillance. This task needs to be efficiently performed in different applications, where various types of information need to be collected from multiple locations. In addition to obtaining potentially valuable information (that can often be time-sensitive), one also needs to ensure that the information can be efficiently transmitted between the nodes in a wireless communication/sensor network. In the simplest static case, the location of sensors (i.e., nodes in a sensor network) is fixed, and the links (edges in a sensor network) are determined by the distance between sensor nodes, that is, two nodes would be connected if they are located within their wireless transmission range. However, in many practical situations, the sensors are installed on moving vehicles (for instance, unmanned air vehicles (UAVs)) that can dynamically move within a specified area of surveillance. Clearly, in this case the location of nodes and edges in a network and the overall network topology can change significantly over time. The task of crucial importance in these settings is to develop optimal strategies for these dynamic sensor networks to operate efficiently in terms of both collecting valuable information and ensuring robust wireless connectivity between sensor nodes.

In terms of collecting information from different locations (sites), one needs to deal with the challenge that the number of sites that need to be visited to gather potentially valuable information is usually much larger than the number of sensors. Under these conditions, one needs to develop efficient schedules for all the moving sensors such that the amount of valuable information collected by the sensors is maximized. A relevant

[1] This chapter is based on the joint publication with A. Veremyev, V. Boginski, D.E. Jeffcoat and S. Uryasev (Kalinchenko et al. (2011))

approach that was previously used by the co-authors to address this challenge dealt with formulating this problem in terms of minimizing the information losses due to the fact that some locations are not under surveillance at certain time moments. In these settings, the information losses can be quantified as both fixed and variable losses, where fixed losses would occur when a given site is simply not under surveillance at some time moment, while variable losses would increase with time depending on how long a site has not been visited by a sensor. Taking into account variable losses of information is often critical in the cases of dealing with strategically important sites that need to be monitored as closely as possible. In addition, the parameters that quantify fixed and variable information losses are usually uncertain, therefore, the uncertainty and risk should be explicitly addressed in the corresponding optimization problems.

The other important challenge that will be addressed in this chapter is ensuring robust connectivity patterns in dynamic sensor networks. These robustness properties are especially important in uncertain and adverse environments in military settings, where uncertain failures of network components (nodes and/or edges) can occur.

The considered robust connectivity characteristics will deal with different parameters of the network. First, the nodes within a network should be connected by paths that are not excessively long, that is, the number of intermediary nodes and edges in the information transmission path should be small enough. Second, each node should be connected to a significant number of other nodes in a network, which would provide the possibility of multiple (backup) transmission paths in the network, since otherwise the network topology would be vulnerable to possible network component failures.

Clearly, the aforementioned robust connectivity properties are satisfied if there are direct links between all pairs of nodes, that is, if the network forms a clique. Cliques are very robust network structures, due to the fact that they can sustain multiple network component failures. Note that any subgraph of a clique is also a clique, which implies that this structure would maintain robust connectivity patterns even if multiple nodes in the network are disabled. However, the practical drawbacks of cliques include the fact that these structures are often overly restrictive and expensive to construct.

To provide a tradeoff between robustness and practical feasibility, certain other network structures that “relax” the definition of a clique can be utilized. The following definitions address these relaxations from different perspectives. Given a graph $G(V, E)$ with a set of vertices (nodes) V and a set of edges E, a k-clique C is a set of vertices in which any two vertices are distance at most k from each other in G Luce (1950). Let $d_G(i, j)$ be the length of a shortest path between vertices i and j in G and $d(G) = \max_{i,j \in V} d_G(i, j)$ be the diameter of G.

Thus, if two vertices $u, v \in V$ belong to a k-clique C, then $d_G(u, v) \leq k$, however this does not imply that $d_{G(C)}(u, v) \leq k$ (that is, other nodes in the shortest path between u and v are not required to belong to the k-clique). This motivated Mokken (Mokken (1979)) to introduce the concept of a k-club. A k-club is a subset of vertices $D \subseteq V$ such that the diameter of induced subgraph $G(D)$ is at most k (that is, there exists a path of length at most k connecting any pair of nodes within a k-club, where all the nodes in this path also belong to this k-club). Also, $\tilde{V} \subseteq V$ is said to be a k-plex if the degree of every vertex in the induced subgraph $G(\tilde{V})$ is at least $|\tilde{V}| - k$ (Seidman and Foster (1978)). A comprehensive study of the maximum k-plex problem is presented in a recent work by Balasundaram et al. (2010).

In this chapter, we utilize these concepts to develop rigorous mathematical programming formulations to model robust connectivity structures in dynamic sensor networks. Moreover, these formulations will also take into account various uncertain parameters by introducing quantitative risk measures that minimize or restrict information losses. Overall, we will develop optimal “schedules” for sensor movements that will take into account both the uncertain losses of information and the robust connectivity between the nodes that would allow one to efficiently exchange the collected information.

3.1 Multi-Sensor Scheduling Problems: General Deterministic Setup

This section introduces a preliminary mathematical framework for dynamic multi-sensor scheduling problems. The simplest deterministic one-sensor version of this problem was introduced in Yavuz and Jeffcoat (2007). The one-sensor scheduling problem was then extended and generalized to more realistic cases of multi-sensor scheduling problems, including the setups in uncertain environments in Boyko et al. (2011). In the subsequent sections of this chapter, this setup will be further extended to incorporate robust connectivity issues into the considered dynamic sensor network models.

To facilitate further discussion, we first introduce the following mathematical notations that will be used throughout this chapter. Assume that there are m sensors that can move within a specified area of surveillance, and there are n sites that need to be observed at every discrete time moment $t = 1, \dots, T$. One can initially assume that a sensor can observe only one site at one point of time and can immediately switch to another site at the next time moment. Since m is usually significantly smaller than n, there will be "breaches" in surveillance that can cause losses of potentially valuable information.

A possible objective that arises in practical situations is to build a strategy that optimizes a potential loss that is associated with not observing certain sites at some time moments.

3.1.1 Formulation with Binary Variables

One can introduce binary decision variables

$$x_{i,t} = \begin{cases} 1, & \text{if } i\text{-th site is observed at time } t \\ 0, & \text{otherwise} \end{cases} \tag{3–1}$$

and integer variables $y_{i,t}$ that denote the last time site i was visited as of the end of time t, $i = 1, \dots, n,\ \ t = 1, \dots, T,\ \ m < n$.

One can then associate a fixed penalty a_i with each site i and a variable penalty b_i of information loss. If a sensor is away from site i at time point t, the fixed penalty a_i is incurred. Moreover, the variable penalty b_i is proportional to the time interval when the site is not observed. We assume that the variable penalty rate can be dynamic; therefore, the values of b_i may be different at each time interval. Thus the loss at time t associated with site i is

$$a_i(1 - x_{i,t}) + b_{i,t}(t - y_{i,t}) \tag{3–2}$$

In the considered setup, we want to minimize the maximum penalty over all time points t and sites i

$$\max_{i,t}\{a_i(1 - x_{i,t}) + b_{i,t}(t - y_{i,t})\} \tag{3–3}$$

Furthermore, $x_{i,t}$ and $y_{i,t}$ are related via the following set of constraints. No more than m sensors are used at each time point; therefore

$$\sum_{i=1}^{n} x_{i,t} \leq m, \ \forall t = 1, \dots, T \tag{3–4}$$

Time $y_{i,t}$ is equal to the time when the site i was last visited by a sensor by time t. This condition is set by the following constraints:

$$0 \leq y_{i,t} - y_{i,t-1} \leq t x_{i,t}, \ \forall i = 1, \dots, n, \ \forall t = 1, \dots, T \tag{3–5}$$

$$t x_{i,t} \leq y_{i,t} \leq t, \ \forall i = 1, \dots, n, \ \forall t = 1, \dots, T \tag{3–6}$$

Further, using an extra variable C and standard linearization techniques, we can formulate the multi-sensor scheduling optimization problem in the deterministic setup as the following mixed integer linear program:

$$\min C \tag{3–7}$$

$$\text{s.t.} \quad C \geq a_i(1 - x_{i,t}) + b_{i,t}(t - y_{i,t}), \ \forall i = 1, \dots, n, \ \forall t = 1, \dots, T \tag{3–8}$$

$$\sum_{i=1}^{n} x_{i,t} \leq m, \ \forall t = 1, \dots, T \tag{3–9}$$

$$0 \le y_{i,t} - y_{i,t-1} \le tx_{i,t},\ \forall i = 1, \dots, n,\ \forall t = 1, \dots, T \tag{3–10}$$

$$tx_{i,t} \le y_{i,t} \le t,\ \forall i = 1, \dots, n,\ \forall t = 1, \dots, T \tag{3–11}$$

$$y_{i,0} = 0,\ \forall i = 1, \dots, n \tag{3–12}$$

$$x_{i,t} \in \{0, 1\},\ \forall i = 1, \dots, n,\ \forall t = 1, \dots, T \tag{3–13}$$

$$y_{i,t} \in R, \forall i = 1, \dots, n,\ \forall t = 0, \dots, T \tag{3–14}$$

We allowed relaxation (3–14) of variables $y_{i,t}$ to the space of real numbers, because the constraints (3–5) and (3–6) enforce the feasible values of variables $y_{i,t}$ to be integer.

3.1.2 Cardinality Formulation

Lemma. Constraint (3–13) is equivalent to the following combination of two constraints:

$$0 \le x_{i,t} \le 1\ \forall i = 1, \dots, n,\ \forall t = 1, \dots, T \tag{3–15}$$

$$\text{card}(\mathbf{x}_t) \le \sum_{i=1}^{n} x_{i,t}\ \forall t = 1, \dots, T \tag{3–16}$$

where $\mathbf{x}_t = (x_{1,t}, \dots, x_{n,t})^T$, and card$(\mathbf{x}_t)$ denotes the cardinality function for the vector $\mathbf{x}_t$. By definition, card$(\mathbf{x}_t)$ equals the number of non-zero elements in the input vector $\mathbf{x}_t$.

Proof. Assume the matrix $(x_{i,t})$ satisfies constraint (3–13). Obviously, it then satisfies (3–15). At the same time, for every t, sum of all elements is equal to the number of values 1 in it. And these are the only non-zero elements in it. Therefore, constraint (3–16) is also satisfied.

Now assume the matrix $(x_{i,t})$ does not satisfy constraint (3–13). Thus there is a pair (i_δ, t_δ), for which $x_{i_\delta,t_\delta} = \delta$ and $\delta \ne 0$ and $\delta \ne 1$. If $\delta < 0$ or $\delta > 1$, then constraint (3–15) is violated. Thus, for all pairs (i, t), $0 \le x_{i,t} \le 1$, and $0 < \delta < 1$. Therefore, for all pairs (i, t), card$(x_{i,t}) \ge x_{i,t}$, and card$(\delta) > \delta$. Taking into account that card$(\mathbf{x}_t) = \sum_i$ card$(x_{i,t})$ we conclude that (3–16) is violated. $\diamond$

Now we can write alternative, cardinality formulation for the general deterministic sensor-scheduling problem.

$$\min C \tag{3–17}$$

$$\text{s.t.} \quad C \geq a_i(1 - x_{i,t}) + b_{i,t}(t - y_{i,t}),\ \forall i = 1, \dots, n,\ \forall t = 1, \dots, T \tag{3–18}$$

$$\sum_{i=1}^{n} x_{i,t} \leq m,\ \forall t = 1, \dots, T \tag{3–19}$$

$$0 \leq y_{i,t} - y_{i,t-1} \leq t x_{i,t},\ \forall i = 1, \dots, n,\ \forall t = 1, \dots, T \tag{3–20}$$

$$t x_{i,t} \leq y_{i,t} \leq t,\ \forall i = 1, \dots, n,\ \forall t = 1, \dots, T \tag{3–21}$$

$$y_{i,0} = 0,\ \forall i = 1, \dots, n \tag{3–22}$$

$$0 \leq x_{i,t} \leq 1,\ \forall i = 1, \dots, n,\ \forall t = 1, \dots, T \tag{3–23}$$

$$\text{card}(\mathbf{x}_t) \leq \sum_{i=1}^{n} x_{i,t},\ \forall t = 1, \dots, T \tag{3–24}$$

$$y_{i,t} \in R, \forall i = 1, \dots, n,\ \forall t = 0, \dots, T \tag{3–25}$$

Although the two formulations are equivalent, some optimization solvers, such as Portfolio Safeguard (that will be mentioned later in this chapter), can provide a near-optimal solution faster if the formulation with cardinality constraints is used instead of the one with boolean variables, which may be important in time-critical systems in military settings.

3.2 Quantitative Risk Measures in Uncertain Environments: Conditional Value-at-Risk

To facilitate further discussion on the formulations of the aforementioned problems under uncertainty, in this section we briefly review basic definitions and facts related to the Conditional Value-at-Risk concept.

Conditional Value-at-Risk (CVaR) Rockafellar and Uryasev (2000, 2002); Sarykalin et al. (2008) is a quantitative risk measure that will be used in the models developed in the next section, which will take into account the presence of uncertain parameters. CVaR is closely related to a well-known quantitative risk measure referred to as

Value-at-Risk (VaR). By definition, with respect to a specified probability level α (in many applications the value of α is set rather high, e.g. 95%), the α-VaR is the lowest amount η_α such that with probability α, the loss will not exceed η_α, whereas for continuous distributions the α-CVaR is the conditional expectation of losses above that amount η_α. As it can be seen, CVaR is a more conservative risk measure than VaR, which means that minimizing or restricting CVaR in optimization problems provides more robust solutions with respect to the risk of high losses (Figure 3-1).

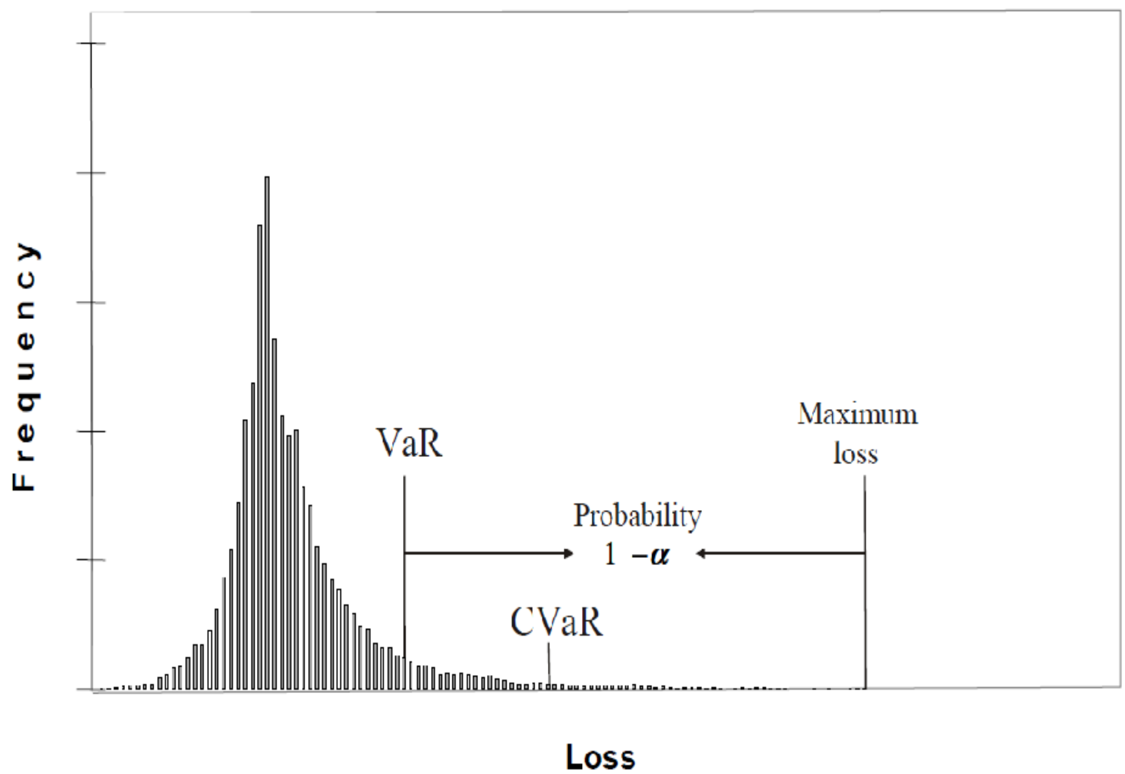

Figure 3-1. Graphical representation of VaR and CVaR.

Formally, α-CVaR for continuous distributions can be expressed as

$$\text{CVaR}_\alpha(\mathbf{x}) = (1-\alpha)^{-1} \int_{L(\mathbf{x},\mathbf{w}) \geq \eta_\alpha(\mathbf{x})} L(\mathbf{x},\mathbf{w})\, p(\mathbf{w})\, d\mathbf{w} \tag{3–26}$$

where $L(\mathbf{x}, \mathbf{w})$ is the random loss (penalty) variable driven by decision vector $\mathbf{x}$ and having a distribution in R induced by that of the vector of uncertain parameters $\mathbf{w}$.

CVaR is defined in a similar way for discrete or mixed distributions. The reader can find the formal definition of CVaR for general case in Rockafellar and Uryasev (2002); Sarykalin et al. (2008).

It has been shown in Rockafellar and Uryasev (2000) that minimizing (3–26) is equivalent to minimizing the function

$$F_\alpha(\mathbf{x},\eta) \;=\; \eta + (1-\alpha)^{-1}\int_{\mathrm{w}\in\mathbb{R}^d}[L(\mathbf{x},\mathbf{w})-\eta]^+\,p(\mathbf{w})\,d\mathbf{w} \tag{3–27}$$

over $\mathbf{w}$ and η, where $[t]^+ = t$ when $t > 0$ but $[t]^+ = 0$ when $t \leq 0$, and optimal value of the variable η corresponds to the VaR value η_α, introduced above.

3.3 Optimizing the Connectivity of Dynamic Sensor Networks Under Uncertainty

This section extends the previous sensors scheduling problem to a stochastic environment. We use CVaR measure to model and optimize various objectives associated with the risk of loss of information.

In the stochastic formulation, the penalties a_i and $b_{i,t}$ are random. We generate S discrete scenarios, which approximate implied joint distribution. Thus, every scenario consists of two arrays: one-dimensional $\{a_i\}^s$ and two-dimensional $\{b_{i,t}\}^s$.

Now, consider the term of the loss function corresponding to the site i, time t, and scenario s:

$$L^s(x,y;i,t) = a_i^s(1-x_{i,t}) + b_{i,t}^s(t-y_{i,t})$$

Under uncertainty, it is often more important to mitigate the biggest possible losses, rather than the average damage. Following this idea, we take $(1-\alpha)$ biggest penalties, and minimize average penalty over all i, t and s. This objective function is exactly the conditional value-at-risk.

We now have the following class of optimization problems:

$$\min_{x,y} CVaR_\alpha\{L(x,y;i,t)\} \tag{3–28}$$

This class has one extreme case: $\alpha = 1$, when the problem becomes equivalent to minimizing maximum possible penalty over all scenarios, locations and time points:

$$\min_{x,y} \max_{i,t,s}(a_i^s(1 - x_{i,t}) + b_{i,t}^s(t - y_{i,t})) \tag{3–29}$$

This problem has an equivalent LP formulation:

$$\min C \tag{3–30}$$

$$\text{s.t.} \quad C \geq a_i^s(1 - x_{i,t}) + b_{i,t}^s(t - y_{i,t}) \tag{3–31}$$

$$\forall i = 1, \dots, n, \forall t = 1, \dots, T, \forall s = 1, \dots, S$$

In order to formulate a general CVaR optimization problem in LP terms we have to introduce additional variables $\tau_{i,t}^s$, $s = 1, \dots, S$, $i = 1, \dots, n$, $t = 1, \dots, T$, and η. With these variables the problem of minimizing CVaR will be reduced to the following:

$$\min C \tag{3–32}$$

$$\text{s.t.} \quad C \geq \eta + \frac{1}{(1-\alpha)nST} \sum_{\substack{s=1,\dots,S \\ i=1,\dots,n \\ t=1,\dots,T}} \tau_{i,t}^s \tag{3–33}$$

$$\tau_{i,t}^s \geq a_i^s(1 - x_{i,t}) + b_{i,t}^s(t - y_{i,t}) - \eta \tag{3–34}$$

$$\forall i = 1, \dots, n, \ t = 1, \dots, T, \ s = 1, \dots, S$$

$$\tau_{i,t}^s \geq 0, \ \forall i = 1, \dots, n, \ t = 1, \dots, T, \ s = 1, \dots, S \tag{3–35}$$

We have discussed various objective functions with objective-specific constraints for sensor scheduling problems in the stochastic environment. In addition to that, every sensor scheduling problem, including those in stochastic environment, must have constraints limiting number of sensors (3–4) and defining variables of the last time of observation (3–5)– (3–6). These constraints are referred to as mandatory constraints for every sensor-scheduling problem.

Further we define a wireless connectivity network $G(V, E)$ on the set of locations V. We interpret it in terms of the 0-1 adjacency matrix $E = \{e_{ij}\}_{i,j=1,\dots,n}$, where each e_{ij} is a

0-1 indicator of wireless signal reachability between nodes i and j, that is, if locations i and j are within direct transmission distance from each other, then they are connected by an edge, and $e_{ij} = 1$ ($e_{ij} = 0$ otherwise). We also define a subnetwork $\tilde{G}$ of $G(V, E)$ containing only those m nodes (locations) that are directly observed by sensors at a particular time moment.

Scheduling of observation often requires sensors to maintain a certain level of wireless connectivity robustness. If an enemy sends a jamming signal that breaks connectivity between a pair of nodes, then the subnetwork $\tilde{G}$ either should stay connected, or at least should maintain unity with probability close to 1. Further, we will utilize several types of network structures that can be applied to ensure that the network satisfies certain robustness constraints.

The most robust network structure is a clique, which implies that each pair of nodes is directly connected by an edge. Obviously, maintaining a clique structure of the subnetwork $\tilde{G}$ at every moment in time is very expensive in terms of penalty, and can be even impossible, if the overall wireless connectivity network is not dense enough. Hence, it is reasonable to utilize appropriate types of clique relaxations to ensure robust network connectivity at every time moment.

One of the considered concepts is a k-plex. By definition, as mentioned above, a k-plex is a subgraph in which every node is connected to at least $m - k$ other nodes in it (where m is the number of nodes in this subgraph). This network configuration ensures that each node is connected to multiple neighbors, which makes it more challenging for an adversary to disconnect the network and isolate the nodes by destroying (jamming) the edges.

Another considered class of network configurations is a k-club. Recall that every pair of nodes in k-club is connected in it through a chain of no more than k arcs (edges). The motivation for studying this type of constraints is based on the fact that if two sensors are connected through a shorter path, it lowers the probability of errors in information transmission through intermediaries, since the number of intermediaries is smaller. Later in the paper, we will specifically use a stronger requirement on the length of these paths. We require that any two nodes are connected either directly by an edge, or through at most one intermediary node, which is often a desired robustness requirement under the conditions when the number of intermediary information transmissions needs to be minimized due to adversarial conditions. Clearly, a 2-club is a structure that satisfies this requirement. In the next subsection, we show that this condition can be incorporated in the considered optimization models.

3.3.1 Ensuring Short Transmission Paths via 2-club Formulation

The general requirement for a subnetwork $\tilde{G}$ to represent a k-club can be formulated as the following set of constraints:

$$\begin{aligned} e_{ij} + \dots \\ &+ \sum_{q=1}^{n} e_{iq} e_{qj} x_{q,t} + \dots \\ &+ \sum_{q=1}^{n} \sum_{l=1}^{n} e_{iq} e_{ql} e_{lj} x_{q,t} x_{l,t} + \dots \\ &+ \sum_{q=1}^{n} \sum_{l=1}^{n} \sum_{p=1}^{n} e_{iq} e_{ql} e_{lp} e_{pj} x_{q,t} x_{l,t} x_{p,t} + \dots \\ &+ \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} \cdots \sum_{i_{k-2}=1}^{n} \sum_{i_{k-1}=1}^{n} e_{ii_1} e_{i_1 i_2} \dots e_{i_{k-2} i_{k-1}} e_{i_{k-1} j} x_{i_1,t} \dots x_{i_{k-1},t} \geq \\ &\geq x_{i,t} + x_{j,t} - 1 \end{aligned} \tag{3–36}$$

where $i = 1, \dots, n-1,\ j = i+1, \dots, n,\ t = 1, \dots, T$. For every k these constraints can be linearized, however, the size of the problem may substantially increase. In this paper, we limit our discussion only to 2-club constraints due to the practical reasons mentioned earlier in this section and due to the fact that the formulation for the case of $k = 2$ will not add too many new entities (no more than $O(n^2)$) to the problem formulation. They require every pair of nodes (i, j) to be connected directly, or through some other node p.

Such type of communication between sensors (i, j) has a concise formulation:

$$e_{ij} + \sum_{p=1}^{n} e_{ip} e_{pj} x_{p,t} \geq x_{i,t} + x_{j,t} - 1 \; \forall i = 1, \dots, n-1, \; \forall j = i+1, \dots, n$$

$$\forall t = 1, \dots, T$$

Here, the left-hand side is always nonnegative. The right-hand side becomes positive only if both locations i and j are observed by sensors, and then it equals 1. According to the 2-club definition, these sensors have to be connected (and exchange information) either directly, or through one other intermediary sensor node. In the first case e_{ij} equals 1. In the second case, the sum $\sum_{p=1}^{n} e_{ip} e_{pj} x_{p,t}$ will also be positive.

It is also important to note that those constraints, for which $e_{ij} = 1$, can be omitted. Thus, a 2-club wireless network configuration can be ensured by the following set of constraints:

$$\sum_{p \in \delta(i) \cap \delta(j)} x_{p,t} \geq x_{i,t} + x_{j,t} - 1$$

$$\forall i = 1, \dots, n-1, \; \forall j = i+1, \dots, n, \; j \notin \delta(i), \; \forall t = 1, \dots, T$$

where $\delta(i)$ and $\delta(j)$ are the sets of neighbors of nodes i and j, respectively.

Below we present the complete general formulation for the dynamic sensor scheduling optimization problem in a stochastic environment with 2-club wireless connectivity constraints.

$$\min C$$

$$\text{s.t.} \quad C \geq \eta + \frac{1}{(1-\alpha)nST} \sum_{\substack{s=1,\dots,S \\ i=1,\dots,n \\ t=1,\dots,T}} \tau_{i,t}^{s}$$

$$\tau_{i,t}^{s} \geq a_i^s (1 - x_{i,t}) + b_{i,t}^{s} (t - y_{i,t}) - \eta$$

$$\forall i = 1, \dots, n, \; t = 1, \dots, T, \; s = 1, \dots, S$$

$$\tau_{i,t}^{s} \geq 0, \; \forall i = 1, \dots, n, \; t = 1, \dots, T, \; s = 1, \dots, S$$

$$\sum_{i=1}^{n} x_{i,t} \leq m, \; \forall t = 1, \dots, T$$

$$0 \leq y_{i,t} - y_{i,t-1} \leq tx_{i,t},\ \forall i = 1, \dots, n,\ \forall t = 1, \dots, T$$

$$tx_{i,t} \leq y_{i,t} \leq t,\ \forall i = 1, \dots, n,\ \forall t = 1, \dots, T$$

$$y_{i,0} = 0,\ \forall i = 1, \dots, n$$

$$\sum_{p \in \delta(i) \cap \delta(j)} x_{p,t} \geq x_{i,t} + x_{j,t} - 1$$

$$\forall i = 1, \dots, n-1,\ \forall j = i+1, \dots, n,\ j \notin \delta(i),\ \forall t = 1, \dots, T$$

$$x_{i,t} \in \{0, 1\},\ \forall i = 1, \dots, n,\ \forall t = 1, \dots, T$$

$$y_{i,t} \in R,\ \forall i = 1, \dots, n,\ \forall t = 0, \dots, T$$

3.3.2 Ensuring Backup Connections via k-plex Formulation

Constraints that require a wireless network to have the k-plex structure, can be defined using a symmetric adjacency matrix $E = \{e_{ij}\}_{i,j=1,\dots,n}$, as defined above. Recall that $\mathbf{x}_t = (x_{1,t}, \dots, x_{n,t})^T$. Consider the vector $\mathbf{z}_t = (z_{1,t}, \dots, z_{n,t}) = E\mathbf{x}_t$. The element $z_{i,t}$ can be interpreted as the number of sensors which have a wireless connection with node i at time t. Thus, the constraint $E\mathbf{x}_t \geq \mathbf{x}_t$ or $(E - I)\mathbf{x}_t \geq 0$ ensures that each sensor node has at least one neighbor, i.e., it is not isolated. If we want each sensor to have at least $(m - k)$ wireless connections (edges) with other sensors, then we should make the constraints more restrictive: $E\mathbf{x}_t \geq (m - k)\mathbf{x}_t$, or

$$(E - (m - k)\, I)\, \mathbf{x}_t \geq 0\ \forall t = 1, \dots, T$$

These restrictions by definition ensure that a subnetwork $\tilde{G}$ is a k-plex.

Below we present the complete general formulation for the dynamic sensor scheduling optimization problem in a stochastic environment with k-plex wireless connectivity constraints.

$$\min C$$

$$\text{s.t.} \qquad C \geq \eta + \frac{1}{(1-\alpha)nST} \sum_{\substack{s=1,\dots,S \\ i=1,\dots,n \\ t=1,\dots,T}} \tau^s_{i,t}$$

$$\tau_{i,t}^s \geq a_i^s(1 - x_{i,t}) + b_{i,t}^s(t - y_{i,t}) - \eta$$

$$\forall i = 1, \dots, n,\ t = 1, \dots, T,\ s = 1, \dots, S$$

$$\tau_{i,t}^s \geq 0,\ \forall i = 1, \dots, n,\ t = 1, \dots, T,\ s = 1, \dots, S$$

$$\sum_{i=1}^{n} x_{i,t} \leq m,\ \forall t = 1, \dots, T$$

$$0 \leq y_{i,t} - y_{i,t-1} \leq t x_{i,t},\ \forall i = 1, \dots, n,\ \forall t = 1, \dots, T$$

$$t x_{i,t} \leq y_{i,t} \leq t,\ \forall i = 1, \dots, n,\ \forall t = 1, \dots, T$$

$$y_{i,0} = 0,\ \forall i = 1, \dots, n$$

$$(E - (m - k)\, I)\, \mathbf{x}_t \geq 0,\ \forall t = 1, \dots, T$$

$$x_{i,t} \in \{0, 1\},\ \forall i = 1, \dots, n,\ \forall t = 1, \dots, T$$

$$y_{i,t} \in R,\ \forall i = 1, \dots, n,\ \forall t = 0, \dots, T$$

3.4 Computational Experiments

Computational experiments on sample problem instances have been performed on Intel Xeon X5355 2.66 GHz CPU with 16GB RAM, using two commercial optimization solvers: ILOG CPLEX 11.2 and AORDA PSG 64 bit (MATLAB 64 bit environment). It should be noted that due to the nature of the considered class of problems, they are computationally challenging even on relatively small networks. Therefore, in many practical situations, finding near-optimal solutions in a reasonable time would be sufficient. The PSG package was used in addition to CPLEX because it has attractive features in terms of coding the optimization problems, and therefore it may be more preferable to use in practical time-critical settings. In particular, in addition to linear and polynomial functions, PSG supports a number of different classes of functions, such as CVaR and cardinality functions. For the purposes of the current case study we defined in PSG the objective using the CVaR function, and we also used cardinality function for the cardinality constraint on $x_{i,t}$ instead of boolean constraint.

Table 3-1. CPLEX Results: Problem with 2-club Constraints

Case		PSG CAR		PSG TANK		CPLEX 26 sec		CPLEX 1 min	
n	m	value	time	value	time	value	time	value	time
10	4	97.96	22.3	100.61	21.6	84.62	26.1	84.62	60.1
10	5	79.29	22.8	83.69	25.6	74.30	26.0	71.57	60.0
10	6	74.79	24.2	71.15	23.6	63.32	26.1	63.32	60.1
10	7	64.82	25.6	61.68	26.5	57.54	26.1	57.54	60.0
10	8	52.27	26.7	51.86	25.2	50.76	26.1	50.21	60.0
11	4	106.19	23.2	105.16	24.2	92.21	26.1	92.21	60.1
11	5	86.61	23.0	85.21	22.4	75.54	26.1	75.54	60.1
11	6	75.08	23.9	76.21	22.5	70.79	26.1	66.32	60.1
11	7	68.43	25.4	69.93	24.0	58.37	26.0	58.37	60.1
11	8	60.44	24.4	61.05	23.8	57.01	26.1	57.01	60.1
12	4	122.32	24.1	124.00	22.8	105.45	26.1	105.45	60.1
12	5	91.25	22.8	98.02	22.6	82.08	26.2	81.96	60.1
12	6	84.69	22.8	81.76	23.0	72.65	26.0	72.65	60.1
12	7	73.44	23.6	76.95	22.1	64.11	26.1	64.11	60.1
12	8	64.78	25.6	61.95	24.0	56.10	26.1	56.10	60.1
13	4	126.37	24.5	119.76	28.9	98.46	26.0	98.46	60.1
13	5	94.48	24.0	104.78	26.5	86.31	26.0	88.20	60.1
13	6	82.46	24.2	83.29	27.2	76.61	26.1	76.61	60.1
13	7	73.97	25.5	74.59	30.8	70.53	26.1	67.93	60.0
13	8	71.57	26.4	69.79	33.9	59.92	26.1	59.92	60.1
14	4	135.75	25.6	139.06	27.3	118.41	26.0	112.74	60.1
14	5	109.75	27.1	114.27	27.2	95.01	26.1	94.87	60.1
14	6	89.58	24.3	93.82	27.2	79.54	26.1	79.54	60.1
14	7	80.70	25.9	80.89	23.3	70.53	26.2	70.37	60.1
14	8	75.31	26.0	76.88	26.6	65.26	26.1	61.67	60.1
15	4	155.67	27.9	145.00	26.7	127.00	26.2	126.82	60.1
15	5	113.18	25.8	115.65	28.8	104.06	26.1	102.34	60.1
15	6	95.51	24.8	99.11	28.1	90.80	26.1	82.74	60.1
15	7	85.96	25.0	86.49	27.8	74.22	26.1	74.22	60.1
15	8	77.81	26.1	76.83	34.4	68.09	26.1	68.22	60.0
			avg		avg		avg		avg
			24.8		26.0		26.1		60.1

For comparison purposes, multiple experiments have been performed. All experiments were divided into two groups: with 2-club connectivity constraints on subnetwork $\tilde{G}$, and k-plex constraints with $k = \left[\frac{m}{2}\right]$. In each of these groups, number of locations $n = 10, 11, 12, 13, 14, 15$ and number of sensors $m = 4, 5, 6, 7, 8$. All problems have CVaR-type objective with $\alpha = 0.9$, deterministic setup (1 scenario), 20

Table 3-2. CPLEX Results: Problem with k-plex Constraints

Case		PSG CAR		PSG TANK		CPLEX 27 sec		CPLEX 1 min	
n	m	value	time	value	time	value	time	value	time
10	4	106.52	25.1	102.94	22.7	85.34	27.1	85.34	60.1
10	5	79.98	23.7	83.48	22.0	72.40	27.1	72.40	60.1
10	6	74.22	25.0	78.25	26.0	63.22	27.0	63.22	60.1
10	7	65.71	27.0	62.27	26.9	56.73	27.0	56.73	60.0
10	8	55.70	25.8	50.91	24.3	50.26	27.0	50.26	60.0
11	4	117.61	23.7	101.28	26.8	87.57	27.0	87.57	60.1
11	5	89.16	24.7	92.77	22.9	75.97	27.1	75.97	60.0
11	6	77.27	24.3	84.08	26.2	68.25	27.1	67.03	60.0
11	7	70.4	26.0	67.24	23.5	58.61	27.1	58.61	60.0
11	8	67.98	25.9	64.84	23.4	53.80	27.1	53.15	60.0
12	4	134.62	31.1	128.60	26.1	109.89	27.1	102.52	60.1
12	5	100.80	24.4	103.70	23.0	80.11	27.1	80.11	60.1
12	6	83.21	24.3	87.56	25.5	70.54	27.1	70.54	60.1
12	7	69.91	25.8	74.99	26.2	63.26	27.1	63.26	60.1
12	8	70.03	27.9	69.28	22.7	56.75	27.1	56.73	60.1
13	4	134.39	32.2	121.72	28.5	103.87	27.1	103.87	60.1
13	5	97.24	25.4	103.89	23.7	84.63	27.1	84.57	60.1
13	6	90.51	25.7	89.40	29.5	77.94	27.1	77.94	60.1
13	7	78.15	26.2	77.34	28.7	68.52	27.1	68.52	60.1
13	8	77.49	26.7	72.61	24.3	63.36	27.1	59.69	60.1
14	4	134.01	30.3	140.12	32.1	119.17	27.1	112.18	60.1
14	5	114.72	26.9	113.61	27.6	90.00	27.1	89.34	60.1
14	6	97.83	27.4	96.71	29.1	78.37	27.1	78.37	60.1
14	7	86.09	26.4	87.00	31.4	70.46	27.1	70.24	60.1
14	8	77.78	27.2	75.75	26.0	62.48	27.0	62.48	60.1
15	4	153.18	34.4	189.41	32.3	136.89	27.1	120.81	60.1
15	5	123.61	28.6	123.84	27.0	110.94	27.1	98.15	60.2
15	6	97.53	27.2	101.72	31.4	83.11	27.0	82.59	60.1
15	7	93.21	27.3	86.48	30.1	74.43	27.1	74.43	60.0
15	8	80.29	28.9	75.70	25.7	67.08	27.1	67.37	60.1
			avg		avg		avg		avg
			26.9		26.5		27.1		60.1

time intervals. The edge density of the considered overall wireless connectivity network was $\rho = 0.8$ (80% pairs of nodes are connected).

We have run PSG using two built-in solvers: CAR and TANK. These solvers took on average 26 seconds to deliver solution over all cases with 2-club constraints, and 27 seconds for the cases with k-plex constraints. After that we run CPLEX on cases

Table 3-3. CPLEX and PSG Results: Stochastic Setup

		CPLEX			PSG CAR		PSG TANK	
type	S	value	time	gap	value	time	value	time
k-plex	10	85.14	300.1	27.0%	96.78	25.3	96.81	29.4
k-plex	20	89.36	300.2	34.6%	95.10	35.3	96.10	37.1
k-plex	50	92.27	300.5	44.4%	110.06	154.7	97.40	273.1
k-plex	100	93.81	301.7	49.7%	104.57	300.6	115.69	300.6
2-club	10	86.92	300.1	30.5%	100.13	229.9	100.13	300.1
2-club	20	84.92	300.1	32.0%	97.55	79.4	97.55	300.1
2-club	50	89.75	300.5	44.3%	104.30	300.1	103.42	300.1
2-club	100	95.87	301.8	50.7%	116.63	300.2	116.23	300.2

with 2-club constraints with time limit 26 seconds, on cases with k-plex constraints with time limit 27 seconds. Then, we additionally run CPLEX on all cases with time limit 1 minute. Computational results are presented in Table 3-1 and Table 3-2, for the cases with 2-club constraints and k-plex constraints, respectively.

The results show that on average the best solution is produced by CPLEX 1 minute run. Values, obtained by CPLEX runs with 26 and 27 seconds limits are by 1.2% and 2.2% greater for 2-club and for k-plex respectively. In most cases solutions obtained by two runs were equal. Therefore, CPLEX obtains solution close to optimal in about less than 30 seconds. PSG TANK solution value is greater than CPLEX 1 minute solution value by 15.8% and 22.4% for 2-club and for k-plex respectively. PSG CAR performs slightly better than another solver, providing the solution values greater than CPLEX 1 minute solution values by 15.0% and 22.0%.

In addition to deterministic setup, we have run the aforementioned optimization problems under uncertainty on several stochastic problem instances with the number of sensors $m = 6$, the number of locations $n = 12$, the CVaR-type objective with $\alpha = 0.9$, $T = 10$ time intervals, for different numbers of scenarios: $S = 10, 20, 50, 100$. As before, the wireless connectivity network edge density was $\rho = 0.8$. The time limit was set to 5 minutes. PSG solvers in most cases provided solution before the time limit was reached

(Table 3-3). However, the quality of solution was worse then provided by CPLEX by 15% on average.

CHAPTER 4
CALIBRATING RISK PREFERENCES WITH GENERALIZED CAPM BASED ON MIXED CVAR DEVIATION

This chapter is based on the joint publication with S. Uryasev and R.T. Rockafellar (Kalinchenko et al. (2012))

The Capital Asset Pricing Model (CAPM, Sharpe (1964), Lintner (1965), Mossin (1966), Treynor (1961), Treynor (1999)) after its foundation in the 1960's became one of the most popular methodologies for estimation of returns of securities and explanation of their combined behavior. This model assumes that all investors want to minimize risk of their investments, and all investors measure risk by the standard deviation of return. The model implies that all optimal portfolios are mixtures of the Market Fund and risk free instrument. The Market Fund is commonly approximated by some stock market index, such as S&P500.

An important practical application of the CAPM model is the possibility to calculate hedged portfolios uncorrelated with the market. To reduce the risk of a portfolio, an investor can include additional securities and hedge market risk. The risk of the portfolio in terms of CAPM model is measured by "beta". The value of beta for every security or portfolio is proportional to the correlation between its return and market return. This follows from the assumption that investors have risk attitudes expressed with the standard deviation (volatility). The hedging is designed to reduce portfolio beta with the idea to protect the portfolio in case of a market downturn. However, beta is just a scaled correlation with the market and there is no guarantee that hedges will cover losses during sharp downturns, because the protection works only on average for the frequently observed market movements. Recent credit crises have shown that hedges have tendencies to perform very poorly when they are most needed in extreme market conditions. The classical hedging procedures based on standard beta set up a defence around the mean of the loss distribution, but fail in the tails. This deficiency has led to multiple attempts to improve the CAPM.

One approach to CAPM improvement is to include additional factors in the model. For example, Kraus and Litzenberger (1976), Friend and Westerfield (1980), and Lim (1989) provide tests for the three-moment CAPM, including co-skewness term. This model accounts for non-symmetrical distribution of returns. Fama and French (1996) added to the asset return linear regression model two additional terms: the difference between the return on a portfolio of small stocks and the return on a portfolio of large stocks, and the difference between the return on a portfolio of high-book-to-market stocks and the return on a portfolio of low-book-to-market stocks. Recently, Barberis and Huang (2008) presented CAPM extension based on prospect theory, which allows to price security's own skewness.

The second approach is to find alternative risk measures, which may more precisely represent risk preferences of investors. For instance, Konno and Yamazaki (1991) applied an $\mathcal{L}^1$ risk model (based on mean absolute deviation) to the portfolio optimization problem with NIKKEI 225 stocks. Their approach led to linear programming instead of quadratic programming in the classical Markowitz's model, but computational results weren't significantly better. Further research has been focused on risk measures more correctly accounting for losses. For example, Estrada (2004) applied downside semideviation-based CAPM for estimating returns of Internet company stocks during the Internet bubble crisis. Downside semideviation calculates only for the losses underperforming the mean of returns. Nevertheless, semideviation, similarly to standard deviation, doesn't pay special attention to extreme losses, associated with heavy tails. Sortino and Forsey (1996) also point out that downside deviation does not provide complete information needed to manage risk.

A much more advanced line of research is considered in papers of Rockafellar, Uryasev and Zabarankin (Rockafellar et al. (2006a), Rockafellar et al. (2006b), Rockafellar et al. (2007)). The assumption here is that there are different groups of investors having different risk preferences. The generalized Capital Asset Pricing

Model (GCAPM) proposes that there is a collection of deviation measures, representing risk preferences of the corresponding groups of investors. These deviation measures substitute for the standard deviation of the classical theory. With the generalized pricing formula following from GCAPM one can estimate the deviation measure for a specific group of investors from market prices. This is done by considering parametric classes of deviation measures and calibrating parameters of these measure. The GCAPM provides an alternative to the classical CAPM measure of systematic risk, so-called “generalized beta”. Similarly to classical beta, the generalized beta can be used in portfolio optimization for hedging purposes.

We consider the class of so-called mixed CVaR deviations, having several attractive properties. First, different terms in the mixed CVaR deviation give credit to different parts of the distribution. Therefore, by varying parameters (coefficients), one can approximate various structures of risk preferences. In particular, so-called tail-beta can be built which accounts for heavy tail losses (e.g., losses in the top 5% of the tail distribution). Second, mixed CVaR deviation is a “coherent” deviation measure, and it therefore satisfies a number of desired mathematical properties. Third, optimization of problems with mixed CVaR deviation can be done very efficiently. For instance, for discrete distributions, the optimization problems can be reduced to linear programming.

We consider a setup with one group of investors (representative investor). We assume that these investors estimate risks with the mixed CVaR deviations having fixed quantile levels: 50, 75, 85, 95 and 99 percent of the loss distribution. By definition, this mixed CVaR deviation is a weighted combination of average losses exceeding these quantile levels. The weights for CVaRs with the different quantile levels determine a specific instance of the risk measure. The generalized pricing formula and generalized beta for this class of deviation measures are used in this approach. With market option prices the parameters of the deviation measure are calibrated, thus estimating risk preferences of investors.

Several numerical experiments calibrating risk preferences of investors at different time moments were conducted. We have found that the deviation measure, representing investors' risk preferences, has the biggest weight on the $\mathrm{CVaR}_{50\%}$ term, which equals the average loss below median return. On average, about 11% of the weight is assigned to $\mathrm{CVaR}_{85\%}$, $\mathrm{CVaR}_{95\%}$ and $\mathrm{CVaR}_{99\%}$ evaluating heavy-loss scenarios. Experiments also showed that risk preferences tend to change over time reflecting investors' opinions about the state of market.

We are not the first who attempted to extract risk preferences from option prices. It is a common knowledge that option prices convey risk neutral probability distribution. Some studies, such as Ait-Sahalia and Lo (2000), Jackwerth (2000), Bliss and Panigirtzoglou (2002), contain various approaches to extracting risk preferences in the form of utility function by comparing objective (or statistical) probability density function with risk neutral probability density function, estimated from option prices. In our work risk preferences are expressed in the form of deviation measure, thus making it impossible to compare results with previous studies. We believe, however, that a wide range of applicability of generalized CAPM framework make our results being useful in a greater variety of applications in practical finance.

The remaining part of this chapter is structured as follows. Section 2 recalls the necessary background, describes the assumptions of the model, provides the main definitions and statements, and presents the derivation of the generalized pricing formula. Section 3 contains description of the case study. Section 4 presents the results of the case study. The conclusion section provides several ideas for further research that can be performed in this area.

4.1 Description of the Approach

4.1.1 Generalized CAPM Background

In the classical Markovitz portfolio theory (Markowitz (1952)) all investors are mean-variance optimizers. Contrary to the classical approach, consider now a group of investors who form their portfolios by solving optimization problems of the following type:

$$P(\Delta) \qquad \min_{\substack{x_0 r_0 + x^T Er \geq r_0 + \Delta \\ x_0 + x^T e = 1}} \mathcal{D}\left(x_0 r_0 + x^T r\right)$$

where $\mathcal{D}$ is some measure of deviation (not necessary standard deviation), r_0 denotes the risk-free rate of return, r is a column vector of (uncertain) rates of return on available securities, and e is a column-vector of ones. Problem $P(\Delta)$ minimizes deviation of the portfolio return subject to a constraint on its expected return and the budget constraint. Different investors within the considered group may demand different excess return Δ. Unlike classical theory, instead of variance or standard deviation, investors measure risk with their generalized deviation measure $\mathcal{D}$. According to the definition in Rockafellar et al. (2006a), a functional $\mathcal{D} : \mathcal{L}^2 \to [0, \infty]$ is a deviation measure if it satisfies the following axioms:

(D1) $\mathcal{D}(X + C) = \mathcal{D}(X)$ for all X and constants C.

(D2) $\mathcal{D}(0) = 0$ and $\mathcal{D}(\lambda X) = \lambda \mathcal{D}(X)$ for all X and all $\lambda > 0$.

(D3) $\mathcal{D}(X + Y) \leq \mathcal{D}(X) + \mathcal{D}(Y)$ for all X and Y.

(D4) $\mathcal{D}(X) \geq 0$ for all X with $\mathcal{D} > 0$ for nonconstant X.

Similarly to Rockafellar et al. (2006b), we can eliminate x_0, which is equal to $1 - x^T e$:

$$P^0(\Delta) \qquad \min_{x^T(Er - r_0 e) \geq \Delta} \mathcal{D}\left(x^T r\right)$$

A pair (x_0, x) is an optimal solution to $P(\Delta)$ if and only if x is an optimal solution to $P^0(\Delta)$ and $x_0 = 1 - x^T e$. Theorem 1 in Rockafellar et al. (2006c) shows that an optimal solution to $P^0(\Delta)$ exists, if deviation measure $\mathcal{D}$ satisfies the property (D5):

(D5) $\{\, X \,|\, \mathcal{D}(X) \leq C \,\}$ is closed for every constant C.

A deviation measure $\mathcal{D}$ satisfying this property is called lower semicontinuous. For further results we will also require an additional property, called lower range dominance:

(D6) $\mathcal{D}(X) \leq EX - \inf X$ for all X.

In this paper we consider only lower semicontinuous, lower range-dominated deviation measures.

Rockafellar et al. (2006b) show that if a group of investors solves problems $P(\Delta)$, the optimal investment policy is characterized by the Generalized One-Fund Theorem (Theorem 2 in that paper). According to the result, the optimal portfolios have the following general structure:

$$x^{\Delta} = \Delta x^{1} \ , \quad x_0^{\Delta} = 1 - \Delta (x^1)^T e$$

where x_0^{Δ} is the investment in risk free instrument, x^{Δ} is a vector of positions in risky instruments, and (x_0^1, x^1) is an optimal solution to $P(\Delta)$, with $\Delta = 1$. Portfolio (x_0^1, x^1) is called a basic fund. It is important to note that, in full generality, $(x^1)^T Er$ could be positive, negative, or equal 0 (threshold case), although for most situations the positive case should prevail.

According to the same paper, a portfolio $x^{\mathcal{D}}$ is called a master fund of positive (negative) type if $(x^{\mathcal{D}})^T e = 1$ ($(x^{\mathcal{D}})^T e = -1$), and $x^{\mathcal{D}}$ is a solution to $P^0(\Delta^*)$ for some $\Delta^* > 0$. From the definition follows that master fund contains only risky securities, with no investment in risk free security. With this definition, the generalized One-Fund Theorem can be reformulated in terms of the master fund. Below we present its formulation as it was given in Rockafellar et al. (2006b).

Theorem 1 (One-Fund Theorem in Master Fund Form). *Suppose a master fund of positive (negative) type exists, furnished by an $x^{\mathcal{D}}$-portfolio that yields an expected return $r_0 + \Delta^*$ for some $\Delta^* > 0$. Then, for any $\Delta > 0$, the solution for the portfolio problem $P(\Delta)$ is obtained by investing the positive amount Δ^*/Δ (negative amount $-\Delta^*/\Delta$) in the master fund, and the amount $1 - (\Delta^*/\Delta)$ (amount $1 + (\Delta^*/\Delta) > 1$) in the risk free instrument.*

From Theorem 1 follows that for every investor in the considered group, the optimal portfolio can be expressed as a combination of investment in the master fund, and investment in the risk free security.

Rockafellar et al. (2007) extends the framework to the case with multiple groups of investors. Every group of investors i, where $i = 1, \dots, I$, solves the problem $P(\Delta)$ with their own deviation measure $\mathcal{D}_i$. It was shown that there exists a market equilibrium, and optimal policy for every group of investors is defined by the Generalized One-Fund Theorem. In this framework investors from different groups may have different master funds. From now on we assume that a generalized deviation measure represents risk preferences of a given group of investors.

Consider a particular group of investors with risk preferences defined by a generalized deviation $\mathcal{D}$. If their master fund is known, the corresponding Generalized CAPM relations can be formulated. The exact relation depends on the type of the master fund. Let r_M denote the rate of return of the master fund. Then

$$r_M = (x^{\mathcal{D}})^T r = \sum_{j=1}^{n} x_j^{\mathcal{D}} r_j$$

where the random variables r_j stand for rates of return on the securities in the considered economy, $x_j^{\mathcal{D}}$ are the corresponding weights of these securities in the master fund, and $\sum_{j=1}^{n} x_j^{\mathcal{D}} = 1$.

The generalized beta of a security j, replacing the classical beta, is defined as follows:

$$\beta_j = \frac{\text{cov}(-r_j, Q_M^{\mathcal{D}})}{\mathcal{D}(r_M)} \tag{4–1}$$

In this formula $Q_M^{\mathcal{D}}$ denotes the risk identifier corresponding to the master fund, taken from the risk envelope corresponding to the deviation measure $\mathcal{D}$. Examples of risk identifiers for specific deviation measures will be presented in the next subsection.

Rockafellar et al. (2006c) derives optimality conditions for problems of minimizing a generalized deviation of the return on a portfolio. The optimality conditions are applied to characterize three types of master funds. Theorem 5 in that paper, presented below, formulates the optimality conditions in the form of CAPM-like relations.

Theorem 2. Let the deviation $\mathcal{D}$ be finite and continuous.

Case 1. An $x^{\mathcal{D}}$-portfolio with $x_1^{\mathcal{D}} + ... + x_n^{\mathcal{D}} = 1$ is a master fund of positive type, if and only if $Er_M > r_0$ and $Er_j - r_0 = \beta_j (Er_M - r_0)$ for all j.

Case 2. An $x^{\mathcal{D}}$-portfolio with $x_1^{\mathcal{D}} + ... + x_n^{\mathcal{D}} = -1$ is a master fund of negative type, if and only if $Er_M > -r_0$ and $Er_j - r_0 = \beta_j (Er_M + r_0)$ for all j.

Case 3. An $x^{\mathcal{D}}$-portfolio with $x_1^{\mathcal{D}} + ... + x_n^{\mathcal{D}} = 0$ is a master fund of threshold type, if and only if $Er_M > 0$ and $Er_j - r_0 = \beta_j Er_M$ for all j.

From now on we call the conditions specified in the Theorem 2 the Generalized CAPM (GCAPM) relations.

4.1.2 Pricing Formulas in GCAPM

Let $r_j = \zeta_j/\pi_j - 1$, where ζ_j is the payoff or the future price of security j, and π_j is the price of this security today.

Similarly to classical theory, pricing formulas can be derived from the Generalized CAPM relations, as it was done in Sarykalin (2008). The following Lemma presents these pricing formulas both in certainty equivalent form, and risk adjusted form.

Lemma 1.

Case 1. If the master fund is of positive type, then

$$\pi_j = \frac{E\zeta_j}{1 + r_0 + \beta_j \left(Er_M^{\mathcal{D}} - r_0\right)} = \frac{1}{1 + r_0}\left(E\zeta_j + \frac{\text{cov}(\zeta_j, Q_M^{\mathcal{D}})}{\mathcal{D}(r_M^{\mathcal{D}})}\left(Er_M^{\mathcal{D}} - r_0\right)\right)$$

Case 2. If the master fund is of negative type, then

$$\pi_j = \frac{E\zeta_j}{1 + r_0 + \beta_j \left(Er_M^{\mathcal{D}} + r_0\right)} = \frac{1}{1 + r_0}\left(E\zeta_j + \frac{\text{cov}(\zeta_j, Q_M^{\mathcal{D}})}{\mathcal{D}(r_M^{\mathcal{D}})}\left(Er_M^{\mathcal{D}} + r_0\right)\right)$$

Case 3. If the master fund is of threshold type, then

$$\pi_j = \frac{E\zeta_j}{1 + r_0 + \beta_j Er_M^{\mathcal{D}}} = \frac{1}{1 + r_0}\left(E\zeta_j + \frac{\text{cov}(\zeta_j, Q_M^{\mathcal{D}})}{\mathcal{D}(r_M^{\mathcal{D}})} Er_M^{\mathcal{D}}\right)$$

See proof in Appendix.

Rockafellar et al. (2007) proved the existence of equilibrium for multiple groups of investors optimizing their portfolios according to their individual risk preferences, and therefore the pricing formulas in Lemma 1 hold true for all groups of investors.

4.1.3 Mixed CVaR Deviation and Betas

Conditional Value-at-Risk has been studied by various researchers, sometimes under different names (expected shortfall, Tail-VaR). We will use notations from Rockafellar and Uryasev (2002). For more details on stochastic optimization with CVaR-type functions see Uryasev (2000), Rockafellar and Uryasev (2000), Rockafellar and Uryasev (2002), Krokhmal et al. (2002), Krokhmal et al. (2006), Sarykalin et al. (2008).

Suppose random variable X determines some financial outcome, future wealth or return on investment. By definition, Value-at-Risk at level α is the α-quantile of the distribution of $(-X)$:

$$\text{VaR}_\alpha(X) = q_\alpha(-X) = -q_{1-\alpha}(X) = -\inf\{\, z \,|\, F_X(z) > 1-\alpha \,\}$$

where F_X denotes the probability distribution function of random variable X.

Conditional Value-at-Risk for continuous distributions equals the expected loss exceeding VaR:

$$\text{CVaR}_\alpha(X) = -E[X \,|\, X \leq -\text{VaR}_\alpha(X)]$$

This formula underlies the name of CVaR as conditional expectation. For the general case the definition is more complicated, and can be found, for example, in Rockafellar and Uryasev (2000). Conditional Value-at-Risk deviation is defined as follows:

$$\text{CVaR}^\Delta_\alpha(X) = \text{CVaR}_\alpha(X - EX)$$

As follows from Theorem 1 in Rockafellar et al. (2006a), there exists a one-to-one correspondence between lower-semicontinuous, lower range-dominated deviation measures $\mathcal{D}$ and convex positive risk envelopes $\mathcal{Q}$:

$$\mathcal{Q} = \left\{ Q \middle| Q \geq 0,\ EQ = 1,\ EXQ \geq EX - \mathcal{D}(X) \text{ for all } X \right\}$$

$$\mathcal{D}(X) = EX - \inf_{Q \in \mathcal{Q}} EXQ \tag{4–2}$$

The random variable $Q_X \in \mathcal{Q}$, for which $\mathcal{D}(X) = EX - EXQ_X$, is called the risk identifier, associated by $\mathcal{D}$ with X.

For a given X and CVaR deviation, the risk identifier can be viewed as a step function, with a jump at the quantile point:

$$Q_X(\omega) = \frac{1}{1-\alpha} \mathbb{1}\left\{ X(\omega) \leq q_{1-\alpha}(X) \right\} \tag{4–3}$$

where ω denotes an elementary event on the probability space, and $\mathbb{1}\left\{ \mathit{condition} \right\}$ is an indicator function, defined on the same probability space, which equals 1 if *condition* is true, and 0 otherwise. Figure 4-1 illustrates the structure of the CVaR risk identifier, corresponding to some random outcome X. For simplicity, the probability space, assumed in the Figure 4-1, is the space of values of the random variable X.

If the group of investors constructs its master fund by minimizing CVaR deviation, and all r_j are continuously distributed, beta for security j has the following expression, derived in Rockafellar et al. (2006c):

$$\beta_j = \frac{\text{cov}(-r_j, Q_M)}{\text{CVaR}_\alpha^\Delta(r_M)} = \frac{E[Er_j - r_j | r_M \leq -\text{VaR}_\alpha(r_M)]}{E[Er_M - r_M | r_M \leq -\text{VaR}_\alpha(r_M)]} \tag{4–4}$$

Classical beta is a scaled covariance between the security and the market. The new beta focuses on events corresponding to big losses in the master fund. For big α ($\alpha > 0.8$), this expression can be called tail-beta.

The following two theorems lead to the definition of mixed CVaR deviation, which is used for the purpose of this paper.

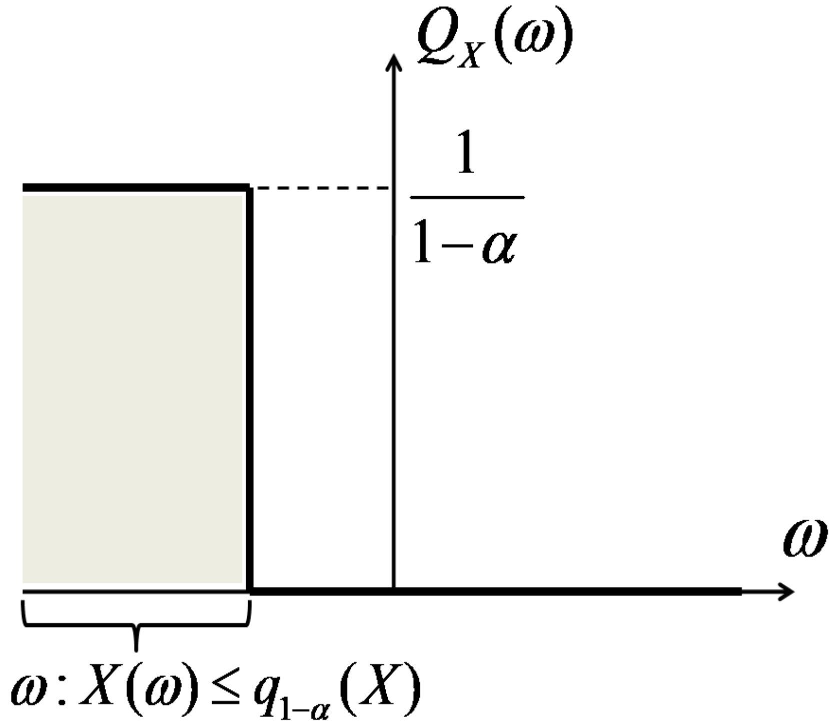

Figure 4-1. CVaR-type Risk Identifier for a Given Outcome Variable X

Theorem 3. Let deviation measure $\mathcal{D}_l$ correspond to risk envelope $\mathcal{Q}_l$ for $l = 1, \dots, L$. If deviation measure $\mathcal{D}$ is a convex combination of the deviation measures $\mathcal{D}_l$:

$$\mathcal{D} = \sum_{l=1}^{L} \lambda_l \mathcal{D}_l \,, \text{ with } \lambda_l \geq 0 \,, \ \sum_{l=1}^{L} \lambda_l = 1$$

then $\mathcal{D}$ corresponds to risk envelope $\mathcal{Q} = \sum_{l=1}^{L} \lambda_l \mathcal{Q}_l$.

See proof in Appendix.

The following theorem presents a formula for the beta corresponding to a deviation measure that is a convex combination of a finite number of deviation measures.

Theorem 4. If the master fund M, corresponding to the deviation measure $\mathcal{D}$, is known, and $\mathcal{D}$ is a convex combination of a finite number of deviation measures $\mathcal{D}_l$, $l = 1, \dots, L$:

$$\mathcal{D} = \lambda_1 \mathcal{D}_1 + \dots + \lambda_L \mathcal{D}_L$$

then

$$\beta_j = \frac{\lambda_1 \text{cov}(-r_j, Q_M^{\mathcal{D}_1}) + ... + \lambda_L \text{cov}(-r_j, Q_M^{\mathcal{D}_L})}{\lambda_1 \mathcal{D}_1(r_M) + ... + \lambda_L \mathcal{D}_L(r_M)}$$

where $Q_M^{\mathcal{D}_l}$ is a risk identifier of master fund return corresponding to deviation measure $\mathcal{D}_l$.

See proof in Appendix.

For a given set of confidence levels $\alpha = (\alpha_1, ..., \alpha_L)$ and coefficients $\lambda = (\lambda_1, ..., \lambda_L)$ such that $\lambda_l \geq 0$ for all $l = 1, ..., L$, and $\sum_{l=1}^{L} \lambda_l = 1$, mixed CVaR deviation $\text{CVaR}^{\Delta}_{\alpha;\lambda}$ is defined in the following way:

$$\text{CVaR}^{\Delta}_{\alpha,\lambda}(X) = \lambda_1 \text{CVaR}^{\Delta}_{\alpha_1}(X) + ... + \lambda_L \text{CVaR}^{\Delta}_{\alpha_L}(X) \tag{4–5}$$

Corollary 1. If $\mathcal{D} = \text{CVaR}^{\Delta}_{\alpha,\lambda}$, where $\alpha = (\alpha_1, ..., \lambda_L)$ and $\lambda = (\lambda_1, ..., \lambda_L)$, and distribution of r_M is continuous, then

$$\beta_j = \frac{\lambda_1 E[Er_j - r_j | r_M \leq -\text{VaR}_{\alpha_1}(r_M)] + ... + \lambda_L E[Er_j - r_j | r_M \leq -\text{VaR}_{\alpha_L}(r_M)]}{\lambda_1 \text{CVaR}^{\Delta}_{\alpha_1}(r_M) + ... + \lambda_L \text{CVaR}^{\Delta}_{\alpha_L}(r_M)} \tag{4–6}$$

See proof in Appendix.

4.1.4 Risk Preferences of a Representative Investor

How can risk preferences of investors be extracted from market prices?

According to GCAPM, risk preferences of a group of investors are represented by a deviation measure. This deviation measure determines the structure of a master fund. For a known deviation measure and a master fund, a risk identifier for the master fund can be specified. If a joint distribution of payoffs for securities is also known, then one can calculate the betas for securities, and then calculate GCAPM prices for these securities. Therefore, according to GCAPM, the deviation measure and the distribution of payoff determine the price for each security. To estimate the deviation measure, having expected returns on securities and market prices, one can find a candidate deviation measure $\mathcal{D}$ for which the GCAPM prices are equal to the market prices.

In this and following sections the paper considers a setup with one group of investors. In other words, all investors evaluate risks of their investments according to the same deviation measure. Therefore, all further results can be referred to as describing a so-called representative investor. From market equilibrium follows that the master fund for a representative investor is known, and, therefore, can be approximated with a market index, such as S&P500.

Alternatively to standard deviation, which measures the magnitude of possible price changes in both directions, Conditional Value-at-Risk deviation measures the average loss for the α worst-case scenarios. We assume that risk preferences can be expressed with a mixed CVaR deviation, defined by formula (4–5), which is a weighted combination of several CVaR deviations with appropriate weights, to capture different parts of the tail of the distribution.

Among the whole variety of securities traded in the market, in addition to the Index fund itself, we consider S&P500 put options with one month to maturity. By construction, put options' prices provide monetary evaluation of the tails of distribution, so they are expected to be perfect candidate to calibrate coefficients in the mixed CVaR deviation.

To estimate the coefficients $\lambda_1, \dots, \lambda_L$ we will use GCAPM formulas, presented in Theorem 2. Let P_K denote the market price of a put option with strike price K and 1 month to maturity, ζ_K denote its (random) monthly return, and $r_K = \frac{\zeta_K}{P_K} - 1$ denote its (random) return in one month. Let r_M be (random) return on the master fund, with its distribution at this moment assumed to be known; r_0 is the return on a risk free security. If market prices are exactly equal to GCAPM prices, and the deviation measure is a mixed CVaR deviation with fixed confidence levels $\alpha_1, \dots, \alpha_L$, then the set of coefficients $\lambda_1, \dots, \lambda_L$ is a solution to the following system of equations:

$$Er_K - r_0 = \beta_K(\lambda)\,(Er_M - r_0)\,, \quad K = K_1, K_2 \dots, K_{J-1}, K_J \tag{4–7}$$

where

$$\beta_K(\lambda) = \frac{\sum_{l=1}^{L} \lambda_l E[Er_K - r_K | r_M \leq -\mathrm{VaR}_{\alpha_l}(r_M)]}{\sum_{l=1}^{L} \lambda_l \mathrm{CVaR}^{\Delta}_{\alpha_l}(r_M)} \tag{4–8}$$

$$\sum_{l=1}^{L} \lambda_l = 1 \tag{4–9}$$

and

$$\lambda_l \geq 0, \quad l = 1, \dots, L \tag{4–10}$$

Equations (4–7) are GCAPM formulas from Theorem 2, applied to market prices P_K of put options with strike prices $K = K_1, \dots, K_J$, and random payoffs ζ_K. Systematic risk measure $\beta(\lambda)$ is expressed through the coefficients λ_l according to Corollary 1.

By multiplying both sides of equation (4–7) by $\sum_{l=1}^{L} \lambda_l \mathrm{CVaR}^{\Delta}_{\alpha_l}(r_M)$ and taking into account (4–8), we get

$$(Er_K - r_0) \sum_{l=1}^{L} \lambda_l \mathrm{CVaR}^{\Delta}_{\alpha_l}(r_M) = (Er_M - r_0) \sum_{l=1}^{L} \lambda_l E[Er_K - r_K | r_M \leq -\mathrm{VaR}_{\alpha_l}(r_M)] \,, \quad K = K_1, \dots, K_J$$

or, equivalently,

$$\sum_{l=1}^{L} \left((Er_K - r_0)\mathrm{CVaR}^{\Delta}_{\alpha_l}(r_M) - (Er_M - r_0) E[Er_K - r_K | r_M \leq -\mathrm{VaR}_{\alpha_l}(r_M)] \right) \lambda_l = 0, \quad K = K_1, \dots, K_J \tag{4–11}$$

If the number of equations (options with different strike prices K) is greater than the number of variables, then system of equations (4–11) may not have a solution. For this reason we replace the equations (4–11) with alternative expressions with error terms e_K:

$$\sum_{l=1}^{L} \left((Er_K - r_0)\mathrm{CVaR}^{\Delta}_{\alpha_l}(r_M) - (Er_M - r_0) E[Er_K - r_K | r_M \leq -\mathrm{VaR}_{\alpha_l}(r_M)] \right) \lambda_l = e_K, \quad K = K_1, \dots, K_J \tag{4–12}$$

We estimate the coefficients $\lambda_1, \dots, \lambda_L$ as the optimal point to the following optimization problem minimizing a norm of vector $(e_{K_1}, \dots, e_{K_J})$:

$$\min_{\lambda_1,\dots,\lambda_L} \|(e_{K_1}, \dots, e_{K_J})\| \tag{4–13}$$

subject to

$$\sum_{l=1}^{L} \left((Er_K - r_0)\mathrm{CVaR}^{\Delta}_{\alpha_l}(r_M) - (Er_M - r_0)E[Er_K - r_K | r_M \leq -\mathrm{VaR}_{\alpha_l}(r_M)]\right)\lambda_l = e_K, \quad K = K_1, \dots, K_J \tag{4–14}$$

$$\lambda_l \geq 0, \quad l = 1, \dots, L\ , \ \sum_{l=1}^{L} \lambda_l = 1 \tag{4–15}$$

In the above formulation $\|\cdot\|$ is some norm. We consider two norms:

$\mathcal{L}^1$-norm:

$$\|(e_{K_1}, \dots, e_{K_J})\|_1 = \frac{1}{J}\sum_{j=1}^{J} |e_{K_j}| \tag{4–16}$$

and $\mathcal{L}^2$-norm:

$$\|(e_{K_1}, \dots, e_{K_J})\|_2 = \sqrt{\frac{1}{J}\sum_{j=1}^{J} e^2_{K_j}}$$

4.2 Case Study Data and Algorithm

We did 153 experiments of estimating risk preferences, each for a separate date (henceforth: date of experiment) starting with 1/22/1998. Dates were chosen with intervals approximately 1 month in such a way that each date is 1 month prior to a next month option expiration date. However, we present detailed analysis for 12 dates with intervals approximately $1/2$ year starting with 12/23/2004. For every experiment we used a set of S&P500 put options with strike prices $K_1, \dots, K_J$, where K_J is a strike price of the at-the-money option (option with strike price closest to the Index value). We define option market price P_K as an average of BID and ASK prices:

$$P_K = \frac{1}{2}\left(P_{\mathrm{ask},K} + P_{\mathrm{bid},K}\right)$$

We chose K_1 as a minimum strike price, for which the following two conditions are satisfied: 1) Starting with the option K_1, prices P_{K_j} are strictly increasing, i.e. $P_{K_{j+1}} > P_{K_j}$; 2) Open interest for all options in the range is greater than 0.

For every experiment we designed a set of scenarios of monthly Index rates of return in the following way. Observing historical values of S&P500 over the period from 1/1/1994 to 10/1/2010, for every trading day s from historical observations we recorded the value $\widetilde{r_I^{(s)}} = \frac{I_{s+21}}{I_s} - 1$, where I_s is the Index value on day s.

We further calculate implied volatility σ of the at-the-money option (the option with strike price K_J), and the value

$$\widehat{\sigma} = \text{standard deviation}\,(\widetilde{r_I^{(s)}})$$

Next, every scenario return was modified as follows:

$$r_I^{(s)} = \frac{\sigma}{\widehat{\sigma}}\left(\widetilde{r_I^{(s)}} - E\widetilde{r_I}\right) + r_0 + \xi\sigma \tag{4–17}$$

where the value for the monthly risk-free rate of return r_0 was selected equal to 0.01%, and $\xi > 0$ is some parameter. The new scenarios will have volatility equal to the volatility σ of the at-the-money options, and expected return $r_0 + \xi\sigma$. In formula (4–17) the value of ξ was chosen such that expected returns on options are negative. We selected $\xi = \frac{1}{3}$. Numerical experiments showed that results are not very sensitive to the selection of the parameter ξ.

Suppose, for modeling purposes, that the investors' preferences are described by a mixed CVaR deviation with confidence levels 50%, 75%, 85%, 95% and 99%:

$$\mathcal{D}(\lambda) = \sum_{l=1}^{L} \lambda_l \text{CVaR}_{\alpha_l}^{\Delta} \tag{4–18}$$

where

$$L = 5\,,\quad \alpha_1 = 99\%\,,\quad \alpha_2 = 95\%\,,\quad \alpha_3 = 85\%\,,\quad \alpha_4 = 75\%\,,\quad \alpha_5 = 50\%$$

Table 4-1. Case Study Data for Selected Dates

Date of experiment Notation	Index value I_0	Lowest strike price K_{min}	Highest strike price K_{max}
12/23/04	1210.13	1120	1210
6/16/05	1210.96	1080	1210
12/22/05	1268.12	1100	1270
6/22/06	1245.60	1150	1245
12/21/06	1418.30	1310	1420
6/21/07	1522.19	1375	1520
12/20/07	1460.12	1255	1460
6/19/08	1342.83	1110	1345
12/18/08	885.28	630	885
6/18/09	918.37	735	920
12/17/09	1096.08	900	1095
6/17/10	1116.04	940	1115

Table 4-2. Case Study Common Data

Decription	Notation	Value
Risk-free monthly interest rate	r_0	0.4125%
Number of terms in mixed CVaR deviation	L	5
Confidence level 1	α_1	99%
Confidence level 2	α_2	95%
Confidence level 3	α_3	85%
Confidence level 4	α_4	75%
Confidence level 5	α_5	50%
Number of scenarios (days)	S	5443

and

$$\lambda_l \geq 0\,, \qquad \sum_{l=1}^{5} \lambda_l = 1 \tag{4–19}$$

The input data for the case study are listed in Table 4-1 and Table 4-2.

Multiple tests demonstrated that the results do not depend significantly on the choice of norm in the optimization problem (4–13). Further in this paper we present results obtained using $\mathcal{L}^1$-norm.

The following steps describe the algorithm, which was used to estimate risk preferences from the option prices.

Step 1. Calculate scenarios indexed by $s = 1, \dots, S$ for payoffs and net returns of put options according to the formula:

$$\zeta_K^{(s)} = \max(0, K - I_0(1 + r_I^{(s)}))\ , \quad r_K^{(s)} = \frac{\xi_K^{(s)}}{P_K} - 1$$

where $K = K_1, \dots, K_J$ are the strike prices, and I_0 is the Index value at time of the experiment.

Step 2. Calculate the following values:

$$E[Er_K - r_K \,|\, r_I \leq -\mathrm{VaR}_{\alpha_l}(r_I)] \ \text{ for all } \ K = K_1, \dots, K_J \ \text{ and } \ l = 1, \dots, L$$

and

$$\mathrm{CVaR}_{\alpha_l}^{\Delta}(r_I) \ \text{ for all } \ l = 1, \dots, L$$

Step 3. Build the design matrix for the constrained regression (4–13)-(4–15) with $r_M = r_I$.

Table 4-3. Deviation Measure Calibration Results

Date of Experiment	$\lambda_{99\%}$	$\lambda_{95\%}$	$\lambda_{85\%}$	$\lambda_{75\%}$	$\lambda_{50\%}$
12/23/2004	0.000	0.020	0.235	0.000	0.745
6/16/2005	0.036	0.016	0.000	0.000	0.948
12/22/2005	0.058	0.000	0.000	0.000	0.942
6/22/2006	0.071	0.033	0.000	0.000	0.895
12/21/2006	0.081	0.000	0.000	0.000	0.919
6/21/2007	0.055	0.040	0.000	0.000	0.905
12/20/2007	0.000	0.041	0.275	0.000	0.684
6/19/2008	0.000	0.055	0.181	0.000	0.765
12/18/2008	0.000	0.000	0.115	0.000	0.885
6/18/2009	0.015	0.014	0.168	0.000	0.803
12/17/2009	0.049	0.001	0.083	0.000	0.868
6/17/2010	0.041	0.048	0.023	0.000	0.889
mean (12 dates)	0.034	0.022	0.090	0.000	0.854
standard deviation (12 dates)	0.030	0.020	0.102	0.000	0.085
mean (153 dates)	0.029	0.029	0.052	0.007	0.883
standard deviation (153 dates)	0.028	0.033	0.077	0.047	0.078

Step 4. Find a set of coefficients λ_l by solving constrained regression (4–13)-(4–15) with $\mathcal{L}^1$ norm, given by equation (4–16). Vector λ gives coefficients in mixed CVaR deviation.

4.3 Case Study Computational Results

Computations were performed on the laptop PC with Intel Core2 Duo CPU P8800 2.66GHz, 4GB RAM and Windows 7, 64-bit. Algorithm, described in previous section, was programmed in MATLAB. Both optimization problems, the constrained regression and CVaR portfolio optimization, on each iteration of the algorithm were solved with AORDA Portfolio Safeguard decision support tool (PSG (2009)). For one date the computational time is around 15 seconds.

The set of coefficients in the mixed CVaR deviation for every date is presented in Table 4-3. This table shows that in all experiments the obtained deviation measure has the biggest weight on $\mathrm{CVaR}_{50\%}$, and smaller weights on either $\mathrm{CVaR}_{85\%}$, $\mathrm{CVaR}_{95\%}$, or $\mathrm{CVaR}_{99\%}$. This can be interpreted as that investors are concerned both with the middle part of the loss distribution, expressed with $\mathrm{CVaR}_{50\%}$, and extreme losses expressed with $\mathrm{CVaR}_{85\%}$, $\mathrm{CVaR}_{95\%}$, or $\mathrm{CVaR}_{99\%}$.

Let us denote by π_K the GCAPM option prices, calculated with pricing formulas in Lemma 1, using calculated mixed CVaR deviation measure and the master fund. We mapped the obtained option prices π_K and the market prices P_K into the implied volatility scale. This mapping is defined by the Black-Scholes formula in implicit form. The graphs of π_K and P_K for 12 dates in the scale of monthly implied volatilities are presented in Figures 4-2, 4-3 and 4-4. All graphs show that the GCAPM prices are close to market prices, except for the graphs for 6/16/2005 and for 12/22/2005.

Figure 4-5 compares dynamics of the value $\eta = 1 - \lambda_{50\%}$ on 153 dates of experiment with S&P500 dynamics. High values of η indicate greater investors' apprehension about potential tail losses and greater inclination to hedge their investments in S&P500. It can be seen that risk preferences were relatively stable until 2008, when the distressed

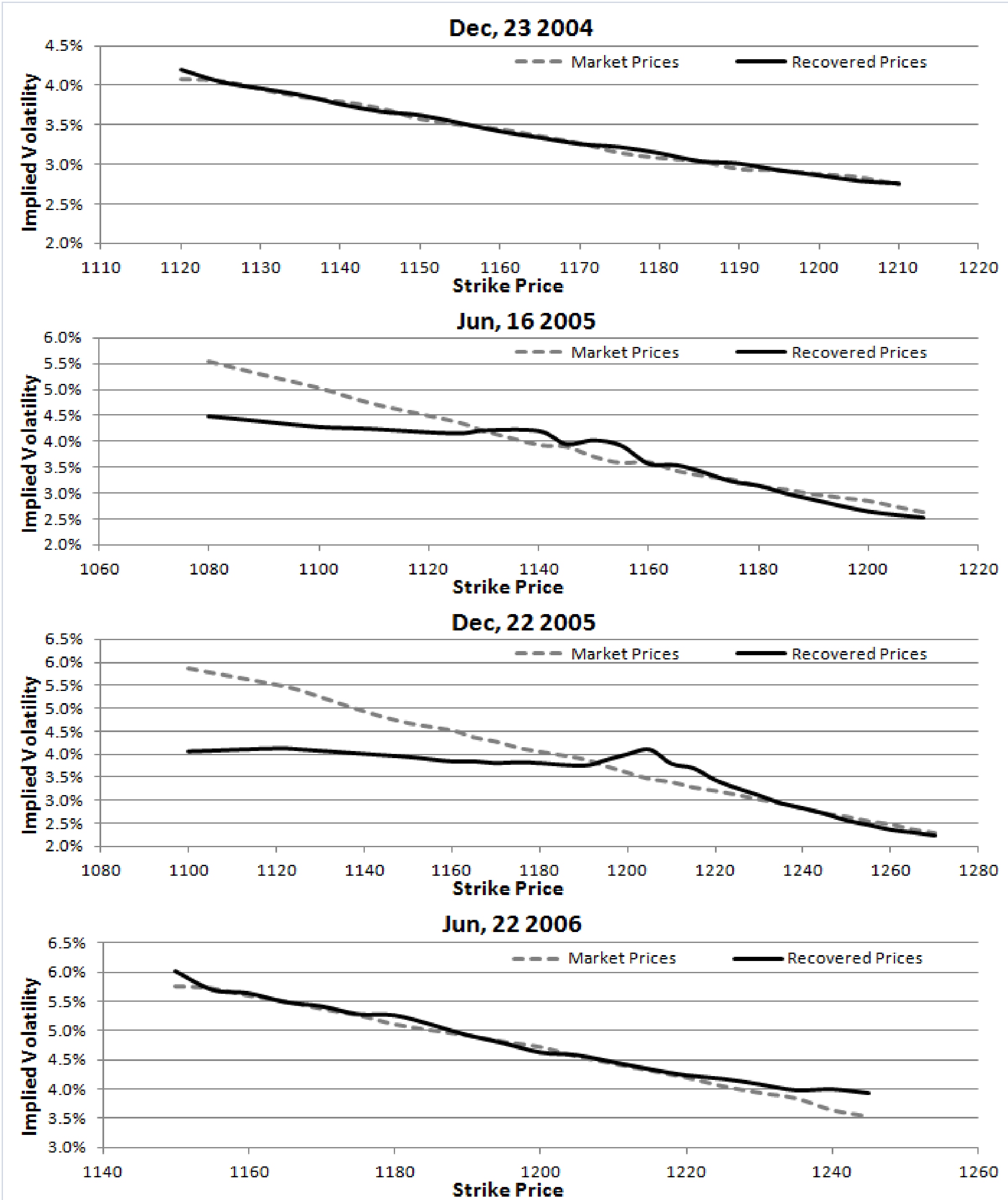

Figure 4-2. Calculated Prices and Market Prices in the Scale of Implied Volatilities
Part 1 out of 3

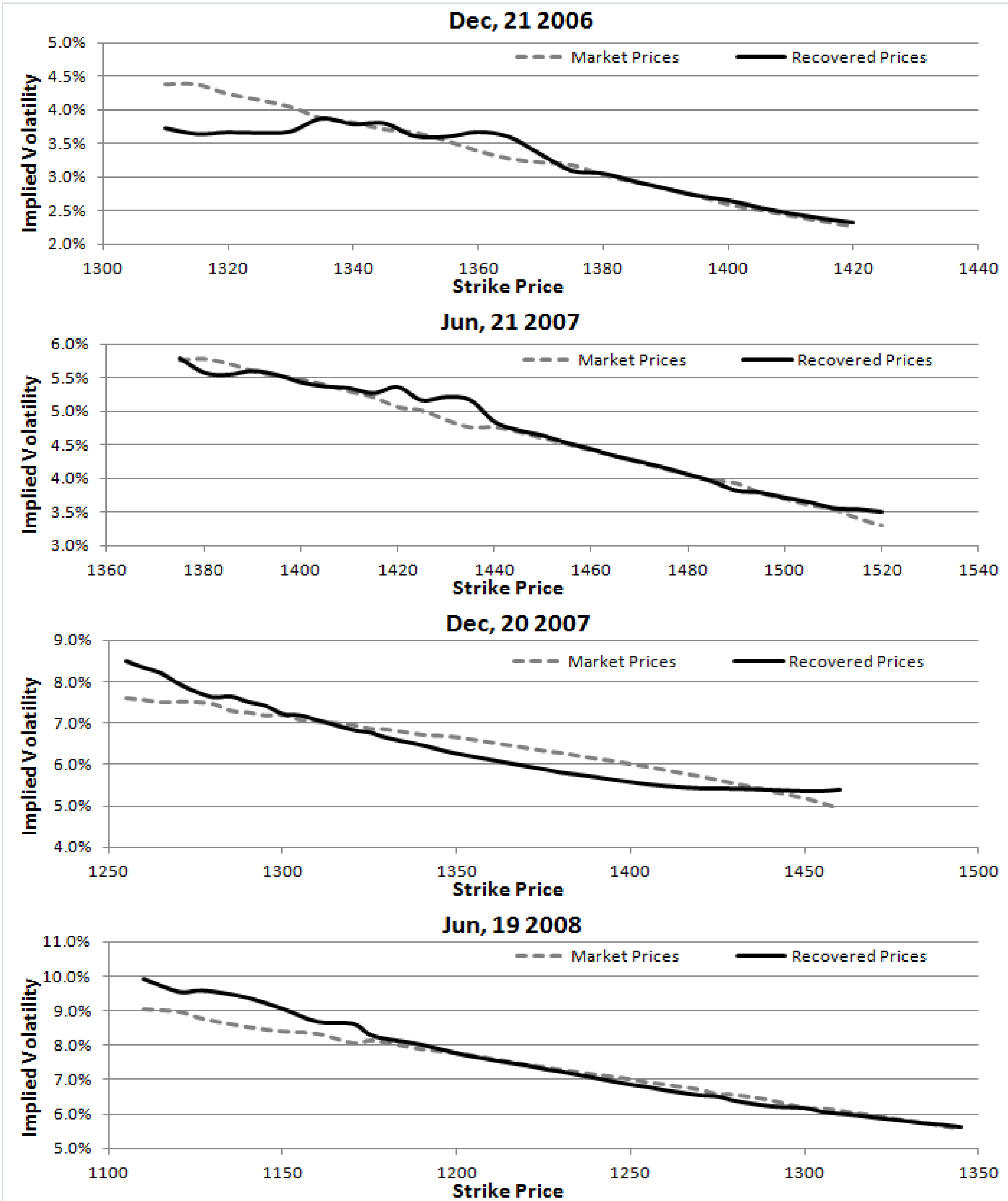

Figure 4-3. Calculated Prices and Market Prices in the Scale of Implied Volatilities Part 2 out of 3

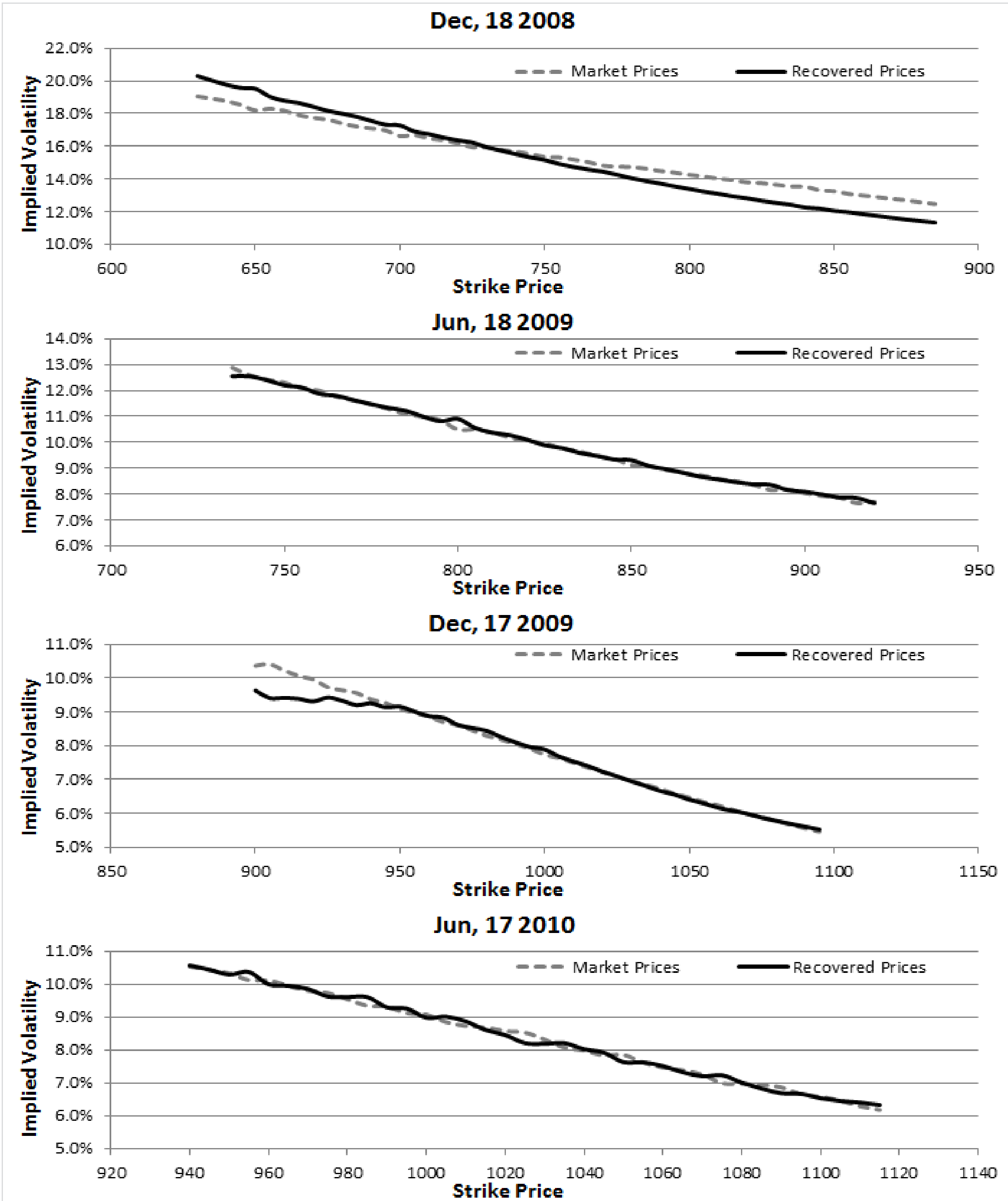

Figure 4-4. Calculated Prices and Market Prices in the Scale of Implied Volatilities Part 3 out of 3

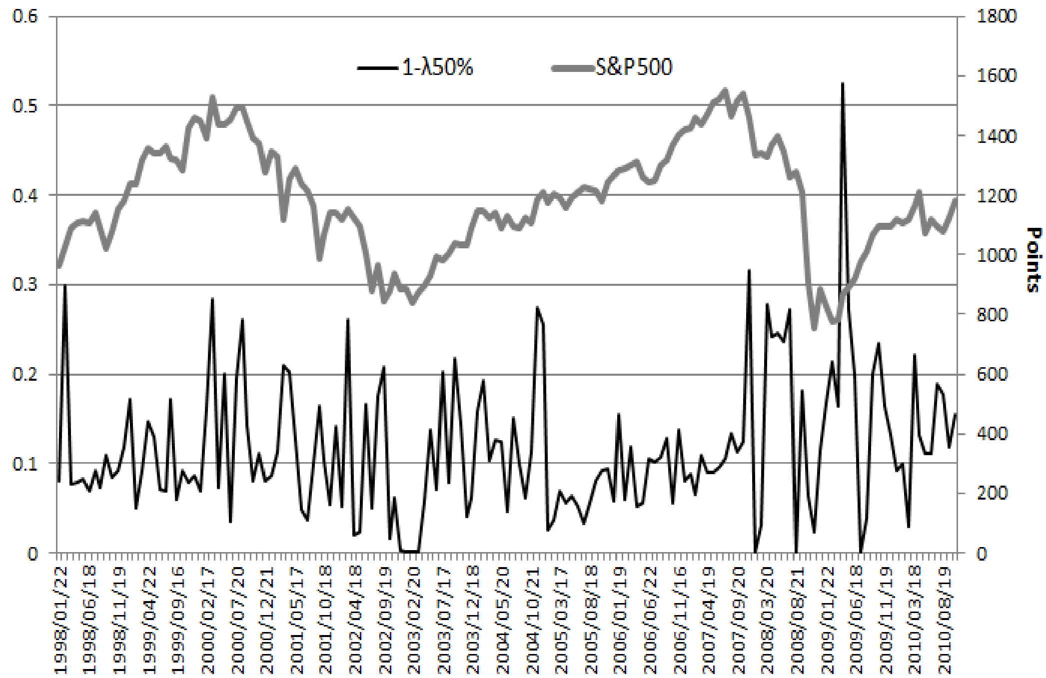

Figure 4-5. S&P500 Value and Risk Aversity Dynamics

period began. It can also be seen that market participants didn't always properly anticipate future market trends. In particular, in December 2008 the value of η was low (0.115), which indicated that market wrongly anticipated that Index reached its bottom and will go up. Nevertheless, 2009 started with further decline in the Index.

CHAPTER 5
CONCLUSIONS

5.1 Dissertation Contribution

In this dissertation we provided an overview of the following classes: generalized deviation measures, risk measures, measures of error, and some of their subclasses. We briefly discussed the motivation for applying these measures in stochastic optimization applications.

We reviewed generalized linear regression models. For the quantile regression we determined the distribution for the residual. This result can be used in applications, which require the distribution of the error term in a quantile factor model to be specified, such as simulation procedures. Interpretation of this distribution as a two-sided exponential distribution makes it possible to estimate various properties of the distribution without applying numerical integration methods. Therefore, implementations of computational models based on this distribution are expected to be highly efficient.

We have defined and extended a class of dynamic sensor scheduling problems, based on conditional value-at-risk, by introducing explicit robust connectivity requirements, specifically, k-club and k-plex constraints, taking into account wireless connectivity requirements for sensors at every time moment. We have also presented computational results for moderate-size instances in both deterministic and stochastic problem setups. Since the size of the stochastic version of the problem is S times larger than for the deterministic version (where S is the number of implied penalty scenarios), solving these stochastic problems is clearly challenging from the computational perspective.

The classes of problems considered in this research are primarily motivated by military applications; however, the developed formulations are general enough so that they can be applied in a variety of settings.

We have described a new technique of expressing risk preferences with generalized deviation measures. We have presented a method for extracting risk preferences from market option prices using these formulas. We have conducted a case study for extracting risk preferences of a representative investor from put option prices.

We extracted risk preferences for 153 dates with 1 month intervals, and expressed them with mixed CVaR deviation. Results demonstrate that investors are concerned both with the middle part of the loss distribution, expressed with $\text{CVaR}_{50\%}$, and extreme losses expressed with $\text{CVaR}_{85\%}$, $\text{CVaR}_{95\%}$, or $\text{CVaR}_{99\%}$. Exact proportions vary, reflecting investors anticipation of high or low returns.

An important application of the theory is that it provides an alternative, more broad view on systematic risk, compared to the classical CAPM based on standard deviation. Similarly to the classical CAPM, we calculated new betas for securities, which measure systematic risk in a different way, capturing tail behavior of a master fund return. These betas can be used for hedging against tail losses, which occur in down market.

Potential applications go beyond identifying risk preferences of considered investors. An investor can express risk attitudes in the form of a deviation measure, and then recalculate betas for securities using this deviation measure. With these betas the investor can build a portfolio hedged according to his risk preferences.

5.2 Future Work

Sensors scheduling problem formulations can be further extended by adding movement network and corresponding constraints as introduced in Boyko et al. (2011), thus modeling the map of possible sensor movements.

It would be interesting to examine investors' perception of risk for different time horizons by estimating risk preferences using market prices of put options with 2, 3, and more months to maturity.

APPENDIX: PROOFS

Below we present proofs of statements formulated in the article. For the reader's convenience, we repeat formulations before every proof.

Lemma 1.

Case 1. If the master fund is of positive type, then

$$\pi_j = \frac{E\zeta_j}{1 + r_0 + \beta_j \left(Er_M^{\mathcal{D}} - r_0\right)} = \frac{1}{1 + r_0}\left(E\zeta_j + \frac{\mathrm{cov}(\zeta_j, Q_M^{\mathcal{D}})}{\mathcal{D}(r_M^{\mathcal{D}})}\left(Er_M^{\mathcal{D}} - r_0\right)\right)$$

Case 2. If the master fund is of negative type, then

$$\pi_j = \frac{E\zeta_j}{1 + r_0 + \beta_j \left(Er_M^{\mathcal{D}} + r_0\right)} = \frac{1}{1 + r_0}\left(E\zeta_j + \frac{\mathrm{cov}(\zeta_j, Q_M^{\mathcal{D}})}{\mathcal{D}(r_M^{\mathcal{D}})}\left(Er_M^{\mathcal{D}} + r_0\right)\right)$$

Case 3. If the master fund is of threshold type, then

$$\pi_j = \frac{E\zeta_j}{1 + r_0 + \beta_j Er_M^{\mathcal{D}}} = \frac{1}{1 + r_0}\left(E\zeta_j + \frac{\mathrm{cov}(\zeta_j, Q_M^{\mathcal{D}})}{\mathcal{D}(r_M^{\mathcal{D}})} Er_M^{\mathcal{D}}\right)$$

Proof of Lemma 1. Proofs for all three cases are similar, so we present the proof only for a master fund of positive type. According to the GCAPM relation specified in Case 1,

$$Er_j - r_0 = \beta_j \left(Er_M^{\mathcal{D}} - r_0\right)$$

Since $r_j = \zeta_j/\pi_j - 1$, then $Er_j = E\zeta_j/\pi_j - 1$, from which we get

$$\frac{E\zeta_j}{\pi_j} - (1 + r_0) = \beta_j \left(Er_M^{\mathcal{D}} - r_0\right) \tag{A–1}$$

This yields the Generalized Capital Asset Pricing Formula in the certainty equivalent form:

$$\pi_j = \frac{E\zeta_j}{1 + r_0 + \beta_j \left(Er_M^{\mathcal{D}} - r_0\right)} \tag{A–2}$$

Using the expression for beta (4–1) we can also write

$$\pi_j = \frac{E\zeta_j}{1 + r_0 + \frac{\mathrm{cov}(-r_j, Q_M^{\mathcal{D}})}{\mathcal{D}(r_M^{\mathcal{D}})}\left(Er_M^{\mathcal{D}} - r_0\right)} \tag{A–3}$$

By multiplying both sides of the equality (A–1) by π_j, we get

$$E\zeta_j - \pi_j(r_0 + 1) = \pi_j\beta_j \left(Er_M^{\mathcal{D}} - r_0\right) \tag{A–4}$$

With expression for beta (4–1) we get

$$\begin{aligned} \pi_j\beta_j &= \pi_j \frac{\text{cov}(-r_j, Q_M^{\mathcal{D}})}{\mathcal{D}(r_M^{\mathcal{D}})} = \frac{\text{cov}(-\pi_j r_j, Q_M^{\mathcal{D}})}{\mathcal{D}(r_M^{\mathcal{D}})} \\ &= \frac{\text{cov}(-\pi_j(1+r_j) + \pi_j, Q_M^{\mathcal{D}})}{\mathcal{D}(r_M^{\mathcal{D}})} = \frac{\text{cov}(-\pi_j(1+r_j), Q_M^{\mathcal{D}})}{\mathcal{D}(r_M^{\mathcal{D}})} + \frac{\text{cov}(\pi_j, Q_M^{\mathcal{D}})}{\mathcal{D}(r_M^{\mathcal{D}})} \end{aligned} \tag{A–5}$$

Here π_j is a constant, consequently the second term in the last sum equals 0. Therefore,

$$\pi_j\beta_j = \frac{\text{cov}(-\pi_j(1+r_j), Q_M^{\mathcal{D}})}{\mathcal{D}(r_M^{\mathcal{D}})}$$

Since $\pi_j(1 + r_j) = \zeta_j$, then

$$\pi_j\beta_j = -\frac{\text{cov}(\zeta_j, Q_M^{\mathcal{D}})}{\mathcal{D}(r_M^{\mathcal{D}})}$$

Substituting expression for $\pi_j\beta_j$ into (A–4) gives:

$$E\zeta_j - \pi_j(r_0 + 1) = -\frac{\text{cov}(\zeta_j, Q_M^{\mathcal{D}})}{\mathcal{D}(r_M^{\mathcal{D}})} \left(Er_M^{\mathcal{D}} - r_0\right)$$

The last equation implies the risk-adjusted form of the pricing formula:

$$\pi_j = \frac{1}{1 + r_0} \left(E\zeta_j + \frac{\text{cov}(\zeta_j, Q_M^{\mathcal{D}})}{\mathcal{D}(r_M^{\mathcal{D}})} \left(Er_M^{\mathcal{D}} - r_0\right) \right) \tag{A–6}$$

◇

Theorem 3. Let deviation measure $\mathcal{D}_l$ correspond to risk envelope $\mathcal{Q}_l$ for $l = 1, \dots, L$. If deviation measure $\mathcal{D}$ is a convex combination of the deviation measures $\mathcal{D}_l$:

$$\mathcal{D} = \sum_{l=1}^{L} \lambda_l \mathcal{D}_l\,, \text{ with } \lambda_l \geq 0\,, \ \sum_{l=1}^{L} \lambda_l = 1$$

then $\mathcal{D}$ corresponds to risk envelope $\mathcal{Q} = \sum_{l=1}^{L} \lambda_l \mathcal{Q}_l$.

Proof of Theorem 3. With formula (4–2) we get:

$$\begin{aligned}\mathcal{D}(X) &= \sum_{l=1}^{L} \lambda_l \mathcal{D}_l(X) = EX - \sum_{l=1}^{L} \lambda_l \inf_{Q \in \mathcal{Q}_l} EXQ = \\ &= EX - \inf_{(Q_1,\dots,Q_L) \in (\mathcal{Q}_1,\dots,\mathcal{Q}_L)} EX \left(\sum_{l=1}^{L} \lambda_l Q_l \right) = EX - \inf_{Q \in \sum_{l=1}^{L} \lambda_l \mathcal{Q}_l} EXQ\end{aligned} \tag{A–7}$$

◇

Theorem 4. If the master fund M, corresponding to the deviation measure $\mathcal{D}$, is known, and $\mathcal{D}$ is a convex combination of a finite number of deviation measures $\mathcal{D}_l$, $l = 1, \dots, L$:

$$\mathcal{D} = \lambda_1 \mathcal{D}_1 + \dots + \lambda_L \mathcal{D}_L$$

then

$$\beta_j = \frac{\lambda_1 \mathrm{cov}(-r_j, Q_M^{\mathcal{D}_1}) + \dots + \lambda_L \mathrm{cov}(-r_j, Q_M^{\mathcal{D}_L})}{\lambda_1 \mathcal{D}_1(r_M) + \dots + \lambda_L \mathcal{D}_L(r_M)}$$

where $Q_M^{\mathcal{D}_l}$ is a risk identifier of master fund return, corresponding to deviation measure $\mathcal{D}_l$.

Proof of Theorem 4. From Theorem 3 follows:

$$\begin{aligned}\beta_j &= \frac{\mathrm{cov}(-r_j, Q_M^{\mathcal{D}})}{\mathcal{D}} = \frac{\mathrm{cov}(-r_j, \lambda_1 Q_M^{\mathcal{D}_1} + \dots + \lambda_L Q_M^{\mathcal{D}_L})}{\lambda_1 \mathcal{D}_1(r_M) + \dots + \lambda_L \mathcal{D}_L(r_M)} = \\ &= \frac{\lambda_1 \mathrm{cov}(-r_j, Q_M^{\mathcal{D}_1}) + \dots + \lambda_L \mathrm{cov}(-r_j, Q_M^{\mathcal{D}_L})}{\lambda_1 \mathcal{D}_1(r_M) + \dots + \lambda_L \mathcal{D}_L(r_M)}\end{aligned} \tag{A–8}$$

Next,

$$\mathrm{cov}(-r_j, Q_M^{\mathcal{D}_l}) = E(Er_j - r_j)(Q_M^{\mathcal{D}_l} - EQ_M^{\mathcal{D}_l}) \tag{A–9}$$

According to the definition of risk envelope, $EQ_M^{\mathcal{D}_l} = 1$. Therefore, from (A–9) we have:

$$\mathrm{cov}(-r_j, Q_M^{\mathcal{D}_l}) = E(Er_j - r_j)(Q_M^{\mathcal{D}_l} - 1) = E(Er_j - r_j)Q_M^{\mathcal{D}_l} - E(Er_j - r_j) = E(Er_j - r_j)Q_M^{\mathcal{D}_l}$$

◇

Corollary 1. If $\mathcal{D} = \text{CVaR}^{\Delta}_{\alpha,\lambda}$, where $\alpha = (\alpha_1, \dots, \lambda_L)$ and $\lambda = (\lambda_1, \dots, \lambda_L)$, and distribution of r_M is continuous, then

$$\beta_j = \frac{\lambda_1 E[Er_j - r_j | r_M \leq -\text{VaR}_{\alpha_1}(r_M)] + \dots + \lambda_L E[Er_j - r_j | r_M \leq -\text{VaR}_{\alpha_L}(r_M)]}{\lambda_1 \text{CVaR}^{\Delta}_{\alpha_1}(r_M) + \dots + \lambda_L \text{CVaR}^{\Delta}_{\alpha_L}(r_M)} \tag{A–10}$$

Proof of Corollary 1. For $\mathcal{D}_l = \text{CVaR}^{\Delta}_{\alpha_l}$, according to (4–3):

$$Q_M^{\mathcal{D}_l}(\omega) = \frac{1}{1-\alpha_l} \mathbb{1}\Big\{ r_M(\omega) \leq -\text{VaR}_{\alpha_l}(r_M) \Big\}$$

Then,

$$\begin{aligned} \text{cov}(-r_j, Q_M^{\mathcal{D}_l}) &= E(Er_j - r_j) \frac{1}{1-\alpha_l} \mathbb{1}\Big\{ r_M(\omega) \leq -\text{VaR}_{\alpha_l}(r_M) \Big\} = \\ &= E[Er_j - r_j | r_M \leq -\text{VaR}_{\alpha_l}(r_M)] \end{aligned} \tag{A–11}$$

Substituting expression for $\text{cov}(-r_j, Q_M^{\mathcal{D}_l})$ and expression for mixed CVaR deviation (4–5) into (A–8) gives:

$$\beta_j = \frac{\lambda_1 E[Er_j - r_j | r_M \leq -\text{VaR}_{\alpha_1}(r_M)] + \dots + \lambda_L E[Er_j - r_j | r_M \leq -\text{VaR}_{\alpha_L}(r_M)]}{\lambda_1 \text{CVaR}^{\Delta}_{\alpha_1}(r_M) + \dots + \lambda_L \text{CVaR}^{\Delta}_{\alpha_L}(r_M)}$$

◇

REFERENCES

Acerbi, Carlo. 2002. Spectral measures of risk: a coherent representation of subjective risk aversion. *Journal of Banking Finance* 26: 1505–1518.

Ait-Sahalia, Yacine, and Andrew Lo. 2000. Nonparametric risk management and implied risk aversion. *Journal of Econometrics* 94: 9–51.

American Optimal Decisions. 2009. *Portfolio Safeguard (PSG)*. URL http://www.aorda.com

Artzner, Philippe, Freddy Delbaen, Jean-Marc Eber, and David Heath. 1999. Coherent measures of risk 9(3): 203–228.

Balasundaram, Balabhaskar, Sergiy Butenko, and Illya V. Hicks. 2010. Clique relaxations in social network analysis: the maximum k-plex problem. *Operations Research* 59(1): 133–142.

Barberis, Nicholas, and Ming Huang. 2008. Stocks as lotteries: The implications of probability weighting for security prices. *American Economic Review* 85(5): 2066–2100.

Basel Committee on Banking Supervision. 2004. *Basel II*. URL http://www.bis.org/bcbs/

Bliss, Robert R., and Nikolaos Panigirtzoglou. 2002. Recovering risk aversion from options. *Journal of Finance* 49(1): 407–446.

Boyko, Nikita, Timofey Turko, Vladimir Boginski, David E. Jeffcoat, Stan Uryasev, Grigoriy Zrazhevsky, and Panos M. Pardalos. 2011. Robust multi-sensor scheduling for sulti-site surveillance. *Journal of Combinatorial Optimization* 22(1): 35–51.

Bradley, Brendan O., and Murad S. Taqqu. 2003. *Financial risk and heavy tails*, chap. 2, 35–104.

Cozzolino, John M., and Michael J. Zahner. 1973. The maximum-entropy distribution of the future market price of a stock. *Operations Research* 21(6): 1200–1211.

Estrada, Javier. 2004. The cost of equity of internet stocks: a downside risk approach. *European Journal of Finance* 239–254.

Fama, Eugene F., and Kenneth R. French. 1996. Multifactor explanations of asset pricing anomalies. *Journal of Finance* 51(1): 55–84.

Friend, Irwin, and Randolph Westerfield. 1980. Co-skewness and capital asset pricing. *Journal of Finance* 35: 897–913.

Grechuk, Bogdan, Anton Molyboha, and Michael Zabarankin. 2009. Maximum entropy principle with general deviation measures. *Mathematics of Operations Research* 34: 445–467.

Jackwerth, Jens Carsten. 2000. Recovering risk aversion from option prices and realized returns. *Review of Financial Studies* 13: 433–467.

Jaynes, Edwin T. 1957. Information theory and statistical mechanics. *The Physical Review* 106(4): 620–630.

Jaynes, Edwin T. 1968. Prior probabilities. *IEEE Transactions on Systems Science and Cybernetics* 4(3): 227–241.

Kalinchenko, Konstantin, Stan Uryasev, and R. Tyrrell Rockafellar. 2012. Calibrating risk preferences with generalized CAPM based on mixed cvar deviation. *Journal of Risk* .

Kalinchenko, Konstantin, Alexander Veremyev, Vladimir Boginski, David E. Jeffcoat, and Stan Uryasev. 2011. Robust connectivity issues in dynamic sensor networks for area surveillance under uncertainty. *Pacific Journal of Optimization* 7(2): 235–248.

Koenker, Roger. 2005. *Quantile Regression*. Econometric Society Monograph Series. Cambridge University Press.

Koenker, Roger, and Gilbert Bassett. 1978. Regression quantiles. *Econometrica* 46(1): 33–50.

Konno, Hiroshi, and Hiroaki Yamazaki. 1991. Mean absolute deviation portfolio optimization model and its application to tokyo stock market. *Management Science* 37: 519–531.

Kraus, Alan, and Robert H. Litzenberger. 1976. Skewness preference and the valuation of risk assets. *Journal of Finance* 31: 1085–1100.

Krokhmal, Pavlo, Jonas Palmquist, and Stan Uryasev. 2002. Portfolio optimization with conditional value-at-risk objective and constraints. *Journal of Risk* 4(2): 43–68.

Krokhmal, Pavlo, R.Tyrrell Rockafellar, and Stan Uryasev. 2006. Risk management and optimization in finance. *Journal of Banking and Finance* 30(2): 315.

Lim, Kian-Guan. 1989. A new test of the three-moment CAPM. *Journal of Financial and Quantitative Analyis* 24: 205–216.

Lintner, John. 1965. The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *Review of Economics and Statistics* 47(1): 13–37.

Luce, R.Duncan. 1950. Connectivity and generalized cliques in sociometric group structure. *Psychometrika* 15: 169–190.

Markowitz, Harry M. 1952. Portfolio selection. *Journal of Finance* 7(1): 77–91.

Mokken, Robert J. 1979. Cliques, clubs and clans. *Quality and Quantity* 13: 161–173.

Mossin, Jan. 1966. Equilibrium in a capital asset market. *Econometrica* 34(4): 768–783.

Pflug, Georg C. 2000. Some remarks on the value-at-risk and the conditional value-at-risk. In *Probabilistic Constrained Optimization: Methodology and Applications*, edited by S. Uryasev, 278–287. Kluwer Academic Publishers.

Rockafellar, R.Tyrrell, and Stan Uryasev. 2000. Optimization of conditional value-at-risk. *Journal of Risk* 2: 21–42.

Rockafellar, R.Tyrrell, and Stan Uryasev. 2002. Conditional value-at-risk for general loss distributions. *Journal of Banking and Finance* 26: 1443–1471.

Rockafellar, R.Tyrrell, Stan Uryasev, and Michael Zabarankin. 2006a. Generalized deviations in risk analysis. *Finance and Stochastics* 10: 51–74.

Rockafellar, R.Tyrrell, Stan Uryasev, and Michael Zabarankin. 2006b. Master funds in portfolio analysis with general deviation measures. *The Journal of Banking and Finance* 30(2): 743–778.

Rockafellar, R.Tyrrell, Stan Uryasev, and Michael Zabarankin. 2006c. Optimality conditions in portfolio analysis with generalized deviation measures. *Mathematical Programming* 108(2-3): 515–540.

Rockafellar, R.Tyrrell, Stan Uryasev, and Michael Zabarankin. 2007. Equilibrium with investors using a diversity of deviation measures. *The Journal of Banking and Finance* 31(11): 3251–3268.

Rockafellar, R.Tyrrell, Stan Uryasev, and Michael Zabarankin. 2008. Risk tuning in generalized linear regression. *Mathematics of Operations Research* 33: 712–729.

Ross, Stephen. 1976. The arbitrage theory of capital asset pricing. *Journal of Economic Theory* 13(3): 341–360.

Sarykalin, Sergey. 2008. *Calibration of general deviation measures from market data*. Ph.D. thesis, University of Florida.

Sarykalin, Sergey, Gaya Serraino, and Stan Uryasev. 2008. Var vs cvar in risk management and optimization. In *Tutorials in Operations Research, INFORMS Tutorial*.

Seidman, Stephen B., and Brian L. Foster. 1978. A graph theoretic generalization of the clique concept. *Journal of Mathematical Sociology* 6: 139–154.

Shannon, Claude Elwood. 1948. A mathematical theory of communication. *Bell System Technical Journal* 27: 379–423, 623–656.

Sharpe, William F. 1964. Capital asset prices: a theory of market equilibrium under conditions of risk. *Journal of Finance* 19: 425–442.

Sortino, Frank A., and Hal J. Forsey. 1996. On the use and misuse of downside risk. *Journal of Portfolio Management* 22(2): 35–42.

Treynor, Jack L. 1961. Market value, time, and risk. Unpublished manuscript.

Treynor, Jack L. 1999. Toward a theory of market value of risky assets. *Asset Pricing and Portfolio Performance: Models, Strategy and Performance Metrics* 15–22.

Uryasev, Stan. 2000. Conditional value-at-risk: optimization algorithms and applications. *Financial Engineering News* 14.

Yavuz, Mesut, and David E. Jeffcoat. 2007. Single sensor scheduling for multi-site surveillance Technical report, Air Force Research Laboratory.

BIOGRAPHICAL SKETCH

Konstantin Kalinchenko was born in 1985 in Protvino, Russia. He received his 5-year Specialist degree in mathematics from Moscow State University University in 2007. Konstantin Kalinchenko worked as an economist and senior economist in 2005-2008 for a leading Russian commercial bank: Sberabank Rossii, where he had an opportunity to take part in developing credit risk measurement methodology and computer methods in control of operations with private securities. In 2008, Konstantin Kalinchenko joined the graduate program in industrial and systems engineering department at the University of Florida. He received his Master of Science degree in industrial and systems engineering from the University of Florida in the spring of 2011, and he received his Ph.D. from the University of Florida in the spring of 2012. Konstantin Kalinchenko also worked as an intern in 2011 for an asset management company State Street Global Advisors (Boston, MA), where he worked on developing quantitative factors based on news parsing solutions for stock ranking. Konstantin Kalinchenko is the author of several scientific papers. He was also a TA in Engineering Economy and Financial Optimization Case Studies classes.