ISE Industrial and Systems Engineering

A Log-Robust Optimization Approach to Portfolio Management

Ban Kawas
Lehigh University

Aurélie Thiele
Lehigh University

Report: 08T-011

A Log-Robust Optimization Approach to Portfolio Management

Ban Kawas* Aurélie Thiele†

December 2007, revised August 2008, December 2008

Abstract

We present a robust optimization approach to portfolio management under uncertainty that builds upon insights gained from the well-known Lognormal model for stock prices, while addressing that model's limitations, in particular, the issue of fat tails being underestimated in the Gaussian framework and the active debate on the correct distribution to use. Our approach, which we call Log-robust in the spirit of the Lognormal model, does not require any probabilistic assumption, and incorporates the randomness on the continuously compounded rates of return by using range forecasts and a budget of uncertainty, thus capturing the decision-maker's degree of risk aversion through a single, intuitive parameter. Our objective is to maximize the worst-case portfolio value (over a set of allowable deviations of the uncertain parameters from their nominal values) at the end of the time horizon in a one-period setting; short sales are not allowed. We formulate the robust problem as a linear programming problem and derive theoretical insights into the worst-case uncertainty and the optimal allocation. We then compare in numerical experiments the Log-robust approach with the traditional robust approach, where range forecasts are applied directly to the stock returns. Our results indicate that the Log-robust approach significantly outperforms the benchmark with respect to 95% or 99% Value-at-Risk. This is because the traditional robust approach leads to portfolios that are far less diversified.

Keywords: robust optimization, portfolio management, convex optimization.

*Department of Industrial and Systems Engineering, Lehigh University, Bethlehem, PA 18015, bhk206@lehigh.edu. Work supported in part by NSF Grant CMMI-0757983.

†P.C. Rossin Assistant Professor, Department of Industrial and Systems Engineering, Lehigh University, Bethlehem, PA 18015, aurelie.thiele@lehigh.edu. Phone: +1-610-758-2903. Work supported in part by NSF Grant CMMI-0757983 and an IBM Faculty Award.

1 Introduction

Portfolio management under uncertainty was pioneered in the 1950s by Markowitz (1959), who first articulated the investor's trade-off between risk and return. The optimal asset allocation, however, is very sensitive to parameter inputs, e.g., mean and covariance, so that small estimation errors can result in strategies that are far from optimal (Chopra and Ziemba (1993)). Goldfarb and Iyengar (2003) have proposed robust optimization approaches to minimize the worst-case variance and similar criteria over ellipsoidal uncertainty sets, which mitigate the impact of estimation errors for these performance measures; their work was later extended to active portfolio management with transaction costs in Erdogan et. al. (2004). The reader is referred to Bertsimas and Thiele (2006) for a tutorial-level introduction to robust optimization. Recently, Value-at-Risk and Conditional Value-at-Risk have emerged as pertinent risk measures in finance, and investors have become more concerned with maximizing the worst-case value of their portfolio than minimizing its standard deviation. This requires the development of new methodologies.

Ben-Tal and Nemirovski (1999) and Bertsimas and Sim (2004) have applied the robust optimization approach not on uncertain parameters but on random variables. In particular, they consider range forecasts at the level of the stock returns, and their goal is to maximize the portfolio's worst-case return, where the worst case is computed over a set of allowable deviations of the stock returns from their nominal values to prevent over-conservatism. This approach is also at the core of the robust financial models developed in Bertsimas and Pachamanova (2008), Pachamanova (2006), Fabozzi et. al. (2007). Numerous studies of stock price behavior, however, (see Hull (2002) and the references therein) suggest that the true drivers of uncertainty are the continuously compounded rates of return, assumed to obey a Gaussian distribution in the famous Lognormal model developed by Black and Scholes (1973). This model gives rise to an elegant mathematical framework and closed-form formulas, for instance for the prices of European options, but neglects the fact that the real distributions have fat tails (Jansen and deVries (1991), Cont (2001), Al Najjab and Thiele (2007)). In that sense, the Lognormal model leads the manager to take more risk than he is willing to accept. Furthermore, the empirical validity of that choice among possible distributions remains open to debate (Fama (1965), Blattberg and Gonedes (1974), Kon (1984), Jansen and deVries (1991), Richardson and Smith (1993), Cont (2001)). In particular, Jansen and deVries (1991) states: "Numerous articles have investigated the distribution of share prices, and find that the returns are fat-tailed. Nevertheless, there is still controversy about the amount of probability mass in the tails, and hence about the most appropriate distribution to use in modeling returns. This controversy has proven hard to resolve." Also, while risk aversion has long been incorporated to portfolio management through the use of utility functions, such functions are difficult to articulate in practice. Robust optimization, however, can capture risk aversion through a single parameter, called the budget of uncertainty, which determines the degree of protection against downside risk the manager requires for his investments.

The decision-maker seeking to protect his portfolio against downside risk needs to find an approach that has the same ease of implementation as the Lognormal model, but which reflects the limited knowledge on the underlying distributions. The purpose of this paper is to provide such an approach for one-period portfolio management, based on robust optimization with polyhedral sets applied to the continuously compounded rates of return. To the best of our knowledge, this is the first time a robust optimization approach has been applied to real-life models of stock price dynamics in portfolio management. We believe this approach gives more relevant results for finance practitioners than the traditional robust approach, while remaining theoretically insightful and numerically tractable.

Contributions. We make the following contributions to the literature.

- We provide a mathematical modeling of uncertainty that builds upon well-established features of stock prices behavior (specifically, the fact that the continuously compounded rates of return are i.i.d.), while addressing limitations of the Lognormal model, where the distribution is assumed to be Gaussian and tail events are underestimated, and of the traditional robust approach, which does not consider the true uncertainty drivers.

- We reformulate the robust one-period portfolio management problem as a linear programming problem, which can be solved efficiently with commercial software, including in large-scale settings with large numbers of stocks.

- We provide insights into the optimal allocation and the worst-case deviations of the uncertain parameters. For instance, we show that when the stock prices are uncorrelated, the amount of money invested in each stock is inversely proportional to the standard deviation of the continuously compounded rate of return, for all the stocks the manager invests in.

- We compare the proposed approach with the traditional robust framework using percentiles of the final portfolio value as performance metrics, and show empirically that the traditional approach leads to much less diversified portfolios, and hence much worse performance, in implementations with real financial data.

Outline. In Section 2 we describe and analyze the portfolio management problem with independent assets. We extend the formulation to the case of correlated stock prices in Section 3. Section 4 contains our numerical experiments. We conclude in Section 5. All proofs are in the appendix.

2 Portfolio Management with Independent Assets

2.1 Generalities

We will use the following notation throughout the paper.

n : the number of stocks,

T : the length of the time horizon,

$S_i(0)$: the initial (known) value of stock i,

$S_i(T)$: the (random) value of stock i at time T,

w_0 : the initial wealth of the investor,

μ_i : the drift of the Lévy process for stock i,

σ_i : the infinitesimal standard deviation of the Lévy process for stock i,

$\tilde{x}_i$: the number of shares invested in stock i,

x_i : the amount of money invested in stock i.

Short sales are not allowed. We start our analysis by assuming all asset prices are independent; this assumption is relaxed in Section 3. Independence between assets arises in particular when managers invest in indices tracking asset classes, such as gold and real estate, rather than stocks. In the traditional Log-normal model (see Hull (2002) for an overview), the random stock price i

at time T, $S_i(T)$, can be described as:

$$\ln \frac{S_i(T)}{S_i(0)} = \left(\mu_i - \frac{\sigma_i^2}{2}\right) T + \sigma_i \sqrt{T} Z_i.$$

where Z_i obeys a standard Gaussian distribution, i.e., $Z_i \sim \mathcal{N}(0,1)$. The portfolio management problem, where the decision-maker seeks to maximize his expected wealth subject to a budget constraint and no short sales, is then formulated as:

$$\begin{array}{ll} \max\limits_{\tilde{x}} & \sum\limits_{i=1}^{n} \tilde{x}_i \mathrm{E}[S_i(T)] \\ \text{s.t.} & \sum\limits_{i=1}^{n} \tilde{x}_i S_i(0) = w_0, \\ & \tilde{x}_i \geq 0, \quad \forall i. \end{array}$$

It is easy to see that the investor will allocate all his wealth to the stock with the highest ratio $\frac{E[S_i(T)]}{S_i(0)}$. To diversify the portfolio in the Lognormal framework, it is then necessary to introduce additional risk constraints, e.g., limiting portfolio variance.

Our goal in this paper is to investigate the optimal asset allocation when the stock price still satisfies Equation (2.1) but the distribution of the Z_i, $i = 1, \ldots, n$ is *not known.* (In particular, it is not necessarily Gaussian.) Instead, we will model Z_i as uncertain parameters with nominal value of zero and known support $[-c, c]$, for all i. Note that all uncertain parameters have the same support, in the spirit of all the random variables in the Lognormal model obeying a standardized Gaussian distribution.

In the remainder of the paper, we will describe the uncertain parameter Z_i, $i = 1, \ldots, n$, as:

$$Z_i = c\,\tilde{z}_i,$$

where $\tilde{z}_i \in [-1, 1]$ represents the *scaled deviation* of Z_i from its nominal value, which is zero. Furthermore, z_i will denote the *absolute value* of the scaled deviation, for all i.

2.2 Problem Setup

To incorporate risk in the formulation, we adopt a worst-case approach where we seek to maximize the worst-case portfolio return over a set of feasible, "realistic" stock returns. The decision-maker has at his disposal the range forecasts $[-c, c]$ for the scaled uncertain parameters Z_i. Furthermore, because the uncertain parameters are assumed *independent*, it is quite unrealistic for many of them to turn out to be equal to their worst-case value; in practice, due to the assumption of independence, some should be higher than their nominal value and some lower, so that a part of the uncertainty will cancel itself out. This motivates the introduction of a budget-of-uncertainty constraint, first presented in Bertsimas and Sim (2004), which bounds the total scaled deviation of the independent, uncertain parameters from their mean (here, zero) by a nonnegative budget denoted Γ:

$$\sum_{i=1}^{n} |\tilde{z}_i| \leq \Gamma, \ |\tilde{z}_i| \leq 1, \ \forall i.$$

If $\Gamma = 0$, all the uncertain parameters are equal to their mean (nominal value). If $\Gamma = n$, the budget-of-uncertainty constraint is redundant with $|\tilde{z}_i| \leq 1$ for all i and the decision-maker will protect the portfolio return against the worst possible value of each stock return. Selecting Γ between these two extremes allows the decision-maker to achieve a trade-off between not protecting the system against any uncertainty and being extremely conservative.

The robust portfolio management problem can then be formulated as:

$$\begin{aligned} \max_{\tilde{x}} \quad & \begin{aligned}[t] \min_{\tilde{z}} \ & \sum_{i=1}^{n} \tilde{x}_i S_i(0) \exp\left[(\mu_i - \frac{\sigma_i^2}{2})T + \sigma_i \sqrt{T} c \tilde{z}_i\right] \\ \text{s.t.} \ & \sum_{i=1}^{n} |\tilde{z}_i| \leq \Gamma, \\ & |\tilde{z}_i| \leq 1 \ \forall i, \end{aligned} \\ \text{s.t.} \quad & \sum_{i=1}^{n} \tilde{x}_i S_i(0) = w_0. \\ & \tilde{x}_i \geq 0 \ \forall i, \end{aligned}$$

or, using that the amount x_i of money invested in stock i at time 0 satisfies: $x_i = S_i(0)\, \tilde{x}_i$ for all

i:

$$\begin{aligned}
\max_{x} \quad & \min_{\tilde{z}} \sum_{i=1}^{n} x_i \exp\left[(\mu_i - \frac{\sigma_i^2}{2})T + \sigma_i\sqrt{T}c\tilde{z}_i\right] \\
& \text{s.t. } \sum_{i=1}^{n} |\tilde{z}_i| \le \Gamma, \\
& \qquad |\tilde{z}_i| \le 1 \ \forall i, \qquad\qquad (1) \\
\text{s.t.} \quad & \sum_{i=1}^{n} x_i = w_0. \\
& x_i \ge 0 \ \forall i.
\end{aligned}$$

Robust optimization addresses the fat-tails issue by specifically planning against the rare (tail) events in the worst-case optimization framework, while these events are under-represented in the traditional Lognormal model, due to the mistakenly low probability estimates.

We observe that the objective in Problem (1) is linear in the asset allocation and nonlinear but convex in the scaled deviations, because short sales are not allowed. In contrast, the traditional robust optimization problem is linear in both the asset allocation and the scaled deviations. Such a modeling approach makes the formulation easier to solve but less relevant in practice, since the true uncertainty driver is inside the exponential term, as mentioned in the literature review. This matters because increasing the deviation of the return $\exp[(\mu_i - \sigma_i^2/2)T + \sigma_i\sqrt{T}c\tilde{z}_i]$, from its nominal value, should not use the same marginal amount of budget if the return is close to its nominal value and if it is already far away from it, due to the nonlinear nature of the function, and especially here where the uncertainty affects the system exponentially.

To ensure that Problem (1) can be solved efficiently using commercial software, our focus will be on rewriting the inner minimization problem as a maximization problem using duality arguments and studying the properties of the resulting formulation.

2.3 Problem Reformulation

We consider the inner minimization problem of Problem (1):

$$\begin{aligned}
\min_{\tilde{z}} \quad & \sum_{i=1}^{n} x_i \exp\left[(\mu_i - \frac{\sigma_i^2}{2})T\right] \exp\left[\sigma_i \sqrt{T} c \tilde{z}_i\right] \\
\text{s.t.} \quad & \sum_{i=1}^{n} |\tilde{z}_i| \leq \Gamma, \\
& |\tilde{z}_i| \leq 1 \qquad \forall i.
\end{aligned} \tag{2}$$

The following lemma allows us to discard the absolute values in Problem (2):

Lemma 2.1 *At optimality, $-1 \leq \tilde{z}_i \leq 0$ for all i and Problem (2) is equivalent to:*

$$\begin{aligned}
\min_{z} \quad & \sum_{i=1}^{n} x_i \exp\left[(\mu_i - \frac{\sigma_i^2}{2})T\right] \exp\left[-\sigma_i \sqrt{T} c z_i\right] \\
s.t. \quad & \sum_{i=1}^{n} z_i \leq \Gamma, \\
& 0 \leq z_i \leq 1, \ \forall i.
\end{aligned} \tag{3}$$

In the remainder of this section, z_i will refer to the absolute value of the scaled deviation and the true worst-case scaled deviation $\tilde{z}_i$ will be negative. Problem (3) is convex; therefore, we study its optimal solution using a Lagrange relaxation approach (see Bertsekas (1999) for a review on nonlinear optimization). For notational convenience, we denote by k_i the constant $\exp\left[(\mu_i - \frac{\sigma_i^2}{2})T\right]$. We introduce the Lagrangian multipliers α, λ_i^0 and λ_i^1 for all i and obtain the unconstrained, convex Lagrange relaxation of Problem (3):

$$\min_{z} \quad \sum_{i=1}^{n} x_i \, k_i \exp(-\sigma_i \sqrt{T} c z_i) + \alpha \left(\sum_{i=1}^{n} z_i - \Gamma \right) - \sum_{i=1}^{n} \lambda_i^0 z_i + \sum_{i=1}^{n} \lambda_i^1 (z_i - 1) \tag{4}$$

Note that strong duality holds because the objective is convex and Slater's condition is satisfied.

Lemma 2.2 (Worst-case deviations)

(i) The optimal deviations z_i (for fixed values of the Lagrange multipliers) are given by, for all i such that $x_i > 0$:

$$z_i = \frac{1}{\sigma_i \sqrt{T} c} \ln \left(\frac{x_i \, k_i \, \sigma_i \sqrt{T} c}{\alpha - \lambda_i^0 + \lambda_i^1} \right). \tag{5}$$

(ii) Specifically, for $x_i > 0$:

- *If $0 < z_i < 1$, then both $\lambda_i^0, \lambda_i^1 = 0$ and Equation (5) becomes:*

$$z_i = \frac{1}{\sigma_i \sqrt{T} c} \left[\ln \left(x_i \, k_i \, \sigma_i \sqrt{T} c \right) - \ln \alpha \right]. \tag{6}$$

- *If $z_i = 0$, then $\lambda_i^1 = 0$ and $\lambda_i^0 \geq 0$ is such that: $x_i \, k_i \, \sigma_i \sqrt{T} c = \alpha - \lambda_i^0$.*

- *If $z_i = 1$, then $\lambda_i^0 = 0$, and $\lambda_i^1 \geq 0$ is such that: $\ln \left(\frac{x_i \, k_i \, \sigma_i \sqrt{T} c}{\alpha + \lambda_i^1} \right) = \sigma_i \sqrt{T} c$, i.e., $\alpha + \lambda_i^1 = x_i \, k_i \, \sigma_i \sqrt{T} c \, \exp(-\sigma_i \sqrt{T} c)$.*

(iii) If $x_i = 0$, then it is optimal to have $z_i = 0$.

Theorem 2.3 shows the robust portfolio management problem (1) can be solved as a linear programming problem.

Theorem 2.3 (Optimal wealth and allocation)

(i) The optimal wealth in the robust portfolio management problem (1) is: $w_0 \exp(F(\Gamma))$, where F is the function defined by:

$$\begin{aligned} F(\Gamma) = \max_{\eta, \chi, \xi} \quad & \sum_{i=1}^{n} \chi_i \ln k_i - \eta \, \Gamma - \sum_{i=1}^{n} \xi_i \\ \text{s.t.} \quad & \eta + \xi_i - \sigma_i \sqrt{T} \, c \, \chi_i \geq 0, \ \forall i, \\ & \sum_{i=1}^{n} \chi_i = 1, \\ & \eta \geq 0, \ \chi_i, \ \xi_i \geq 0, \ \forall i. \end{aligned} \tag{7}$$

(ii) The optimal amount of money invested at time 0 in stock i is $\chi_i \, w_0$, for all i, where the χ_i are found by solving Problem (7).

Remarks.

- The optimal wealth is proportional to the initial amount of money invested, as expected.

- Because Problem (7) is a linear programming problem, it can be solved using the simplex method, yielding a corner point of the feasible set. In particular, the *same* asset allocation (x_i, proportional to χ_i, for all i) will be optimal for a *range* of budgets of uncertainty Γ, which only affect the objective function.

- The worst-case scaled deviations become, for all i: $z_i = \frac{1}{\sigma_i \sqrt{T} \, c} \left[\ln k_i - F(\Gamma) \right]$.

- Using the notations of Section 2.3, we have $\dfrac{\alpha - \lambda_i^0 + \lambda_i^1}{\sigma_i \sqrt{T c}} = w_0 \exp(F(\Gamma))\chi_i$.

- If the investor has additional requirements on the feasible allocation besides the budget constraint and non-negativity, we may still obtain a similar problem formulation using the specific structure of the constraint set. In practice, the easiest way to incorporate additional constraints on the asset allocation is to compute the optimal allocation using the proposed robust optimization approach, and then pick the feasible strategy that is "closest" (for instance, in a least-squares sense) to the theoretical one we have just obtained.

Corollary 2.4 (Scaled deviations) *Assume it is strictly suboptimal to invest in only one stock. Then the scaled deviations for the assets the manager invests in never reach their bounds, i.e.,* $0 < z_i < 1$ *for* i *such that* $x_i > 0$ *at optimality.*

2.4 Structure Of The Optimal Solution

We conclude this section by characterizing the optimal allocation in the Log-robust portfolio management model. We assume that the assets are ranked in decreasing order of their nominal return, i.e., $k_1 < \ldots < k_n$.

Theorem 2.5 (Structure of the optimal allocation) *Assume it is strictly suboptimal to invest in only one stock. At optimality, there exists an index* j *such that the decision-maker invests only in stocks* 1 *to* j *and we have:*

$$x_i = \frac{w_0}{\sigma_i \cdot \left(\sum_{k=1}^{j} \frac{1}{\sigma_k} \right)}. \tag{8}$$

In particular, $x_i\,\sigma_i$ *is constant for all the assets the manager invests in.*

Remarks.

- The optimal allocation does not depend on the scale parameter c, except indirectly through the choice of the diversification parameter j.

- When the manager invests in all assets, Equation (8) is identical to the optimal allocation in the Markowitz mean-variance model when the portfolio variance for independent assets is minimized and the expected return constraint is not binding, but the meaning of the σ_i

is different. In the present approach, the σ_i are the standard deviations of the continuously compounded rates of return. In the Markowitz model, the σ_i traditionally denote the standard deviations of the rates of return over the one period considered. It is natural that the two approaches do not yield the same allocation, since the robust optimization framework does not seek to minimize variance, but instead to maximize worst-case portfolio value.

Diversification. The degree of diversification of the portfolio is determined by the parameter j, the number of stocks the manager invests in. It is easy to see that at optimality, if $\eta > 0$, we have $\eta = \sqrt{T}\,c/(\sum_{k=1}^{j} 1/\sigma_k)$ and $\xi_i = 0$ for all i. We also know that the optimal η is non-increasing with Γ because Problem (7) is piecewise linear, convex, non-increasing in Γ, with slope $-\eta$. Hence, the degree of diversification j of the portfolio is non-decreasing in Γ, as long as $\eta > 0$. When η becomes zero, it becomes optimal to invest in only one stock, specifically the stock with the highest worst-case return. This can be seen mathematically by noting that we have then $\xi_i = \sigma_i\sqrt{T}c\chi_i$ for all i, so that the robust problem becomes:

$$\begin{aligned}
\max_{\chi} \quad & \sum_{i=1}^{n} (\ln k_i - \sigma_i \sqrt{T}\, c)\chi_i \\
\text{s.t.} \quad & \sum_{i=1}^{n} \chi_i = 1, \\
& \chi_i \geq 0, \ \forall i.
\end{aligned}$$

The optimal solution of this linear programming problem is achieved at a corner point of the feasible set, representing an allocation where only one χ_i is non-zero, i.e., only one asset is invested in. The asset selected at optimality is the one with the highest objective coefficient, $\ln k_i - \sum_{i=1}^{n} \sigma_i \sqrt{T}\, c$, which is the asset's worst-case return.

In summary, the log-robust approach yields increased diversification until the manager becomes so risk-averse that he prefers allocating all of his budget to the safest investment.

3 Portfolio Management with Correlated Assets

3.1 Problem Formulation

We now extend the approach described in Section 2 to the case with correlated assets. Equation (2.1), which characterizes the behavior of the stock prices, is replaced by:

$$\ln \frac{S_i(T)}{S_i(0)} = \left(\mu_i - \frac{\sigma_i^2}{2}\right) T + \sqrt{T} Z_i,$$

where the random vector Z is normally distributed with mean $\mathbf{0}$ and covariance matrix $\mathbf{Q}$. We define:

$$\mathbf{Y} = \mathbf{Q}^{-1/2}\mathbf{Z},$$

where $\mathbf{Y} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ and $\mathbf{Q}^{1/2}$ is the square-root of the covariance matrix $\mathbf{Q}$, i.e., the unique symmetric positive definite matrix $\mathbf{S}$ such that $\mathbf{S}^2 = \mathbf{Q}$. In the robust optimization approach, the vector of scaled independent uncertainty drivers $\tilde{\mathbf{y}}$ is related to the vector of (here, non-scaled) deviations $\tilde{\mathbf{z}}$ as follows:

$$\tilde{z}_i = c \sum_{j=1}^{n} Q_{ij}^{1/2}\, \tilde{y}_i,$$

with each component $\tilde{y}_i$ belonging to $[-1, 1]$ so that $Y_i \in [-c, c]$ for all i.

The robust optimization model becomes:

$$\begin{array}{lll}
\max\limits_{x} & \min\limits_{\tilde{y}} & \sum\limits_{i=1}^{n} x_i \exp\left[\left(\mu_i - \sigma_i^2/2\right) T + \sqrt{T} c \left(\sum\limits_{j=1}^{n} Q_{ij}^{1/2}\, \tilde{y}_j\right)\right] \\
& \text{s.t.} & \sum\limits_{j=1}^{n} |\tilde{y}_j| \leq \Gamma, \\
& & |\tilde{y}_j| \leq 1, \ \forall j, \\
\text{s.t.} \ \sum\limits_{i=1}^{n} x_i = w_0, & & \\
\quad\ \ x_i \geq 0, \ \forall i. & &
\end{array} \tag{9}$$

We first need to reformulate the inner minimization problem:

$$\begin{aligned}\min_{\tilde{y}} \quad & \sum_{i=1}^{n} x_i \exp\left[\left(\mu_i - \sigma_i^2/2\right) T\right] \exp\left[\sqrt{T}c\left(\sum_{j=1}^{n} Q_{ij}^{1/2}\, \tilde{y_j}\right)\right] \\ \text{s.t.} \quad & \sum_{j=1}^{n} |\tilde{y_j}| \leq \Gamma, \\ & |\tilde{y_j}| \leq 1, \ \forall j.\end{aligned} \tag{10}$$

as a maximization problem to keep the approach tractable.

Lemma 3.1 *Problem (10) is convex.*

Therefore, we can characterize the optimal solution using a Lagrangean relaxation approach.

3.2 Special Case

In the special case where the coefficients of the square root of the correlation matrix are all non-negative, we observe (using the same argument as in Lemma 2.1) that the minimum of the objective function in Problem (10) is achieved for $\tilde{y_j} \leq 0$. We define $y_j = |\tilde{y_j}|$, so the minimization problem becomes:

$$\begin{aligned}\min_{y} \quad & \sum_{i=1}^{n} x_i \exp\left[\left(\mu_i - \sigma_i^2/2\right) T\right] \exp\left[-\sqrt{T}c\left(\sum_{j=1}^{n} Q_{ij}^{1/2}\, y_j\right)\right] \\ \text{s.t.} \quad & \sum_{j=1}^{n} y_j \leq \Gamma, \\ & 0 \leq y_j \leq 1, \ \forall j.\end{aligned} \tag{11}$$

Lemma 3.2 (Worst-case deviations) *The optimal solution to Problem (11) is given by:*

$$\mathbf{y} = \frac{1}{\sqrt{T}\,c}\,\mathbf{Q}^{-1/2} \ln(\sqrt{T}\,c\,\hat{\mathbf{x}}),$$

where:

$$\hat{x}_i = \frac{x_i \exp\left[\left(\mu_i - \sigma_i^2/2\right) T\right]}{[\mathbf{Q}^{-1/2}(\alpha\,\mathbf{e} - \boldsymbol{\lambda}^0 + \boldsymbol{\lambda}^1)]_i}, \ \forall i.$$

α, $\boldsymbol{\lambda}^0$ and $\boldsymbol{\lambda}^1$ are the Lagrange multipliers associated with the constraints of Problem (11).

Theorem 3.3 shows that the robust portfolio management problem in this special case of correlation can be solved as a linear programming problem; hence, as in the non-correlated case

(Theorem 2.3), the problem can be solved efficiently using commercial software.

Theorem 3.3 (Optimal wealth and allocation)

(i) The optimal wealth in the robust portfolio management problem (9) is: $w_0 \exp(F(\Gamma))$, *where* F *is the function defined by:*

$$\begin{aligned} F(\Gamma) = \max_{\eta, \chi, \xi} \quad & \sum_{i=1}^{n} \chi_i \ln k_i - \eta\, \Gamma - \sum_{i=1}^{n} \xi_i \\ s.t. \quad & \eta + \xi_i - \sqrt{T}\, c \left(\sum_{j=1}^{n} Q_{ij}^{1/2}\, \chi_j \right) \geq 0, \ \forall i, \\ & \sum_{i=1}^{n} \chi_i = 1, \\ & \eta \geq 0, \ \chi_i, \ \xi_i \geq 0, \ \forall i. \end{aligned} \tag{12}$$

(ii) The optimal amount of money invested at time 0 in stock i *is* $\chi_i\, w_0$, *for all* i, *where the* χ_i *are found by solving Problem (12).*

We have seen in Section 2 that, when assets are uncorrelated, $\sigma_i\, \chi_i$ is constant for all the assets i the manager invests in. A similar result, provided in Theorem 3.4, holds in the special correlated case. Because it is an immediate extension to the arguments provided in Section 2, it is stated without proof.

Theorem 3.4 (Structure of the optimal solution) *Let* $S(\Gamma)$ *be the set of indices* j *such that* $\chi_j > 0$ *at optimality. There exists a set* $S_\xi^*(\Gamma)$ *of size* $|S(\Gamma)| - 1$ *such that* $(\chi_j)_{j \in S(\Gamma)}$ *is determined by the set of* $|S(\Gamma)|$ *equations:*

$$\begin{aligned} \sum_{j \in S(\Gamma)} Q_{ij}^{1/2} \chi_j &= \frac{\eta^*(\Gamma)}{\sqrt{T}c}, \ \forall i \in S_\xi^*(\Gamma), \\ \sum_{j \in S(\Gamma)} \chi_j &= 1. \end{aligned}$$

For $j \notin S(\Gamma)$, $\chi_j = 0$.

The robust optimization approach for the correlated case can therefore be interpreted as choosing the sets $S(\Gamma)$ and $S_\xi^*(\Gamma)$ (which depend on each other) appropriately. Once these sets are chosen, the asset allocation is uniquely determined by the equations above.

Diversification. We now comment on how the degree of diversification of the optimal solution varies with Γ. We observe two opposite trends.

Trend 1 The optimal objective of the linear programming problem (7) is convex, non-increasing in Γ (because it can be reformulated as the maximum of linear functions in Γ over all the extreme points of the feasible set), so $\eta^*(\Gamma)$ is non-increasing as well. As a result, the inequalities $\sqrt{T}\, c \sum_{j=1}^{n} Q_{ij}^{1/2}\, \chi_j \geq \eta^*(\Gamma)$ become *less constraining* as Γ increases.

Trend 2 The number of such constraints tends to increase with Γ, because we know from robust optimization theory (see Bertsimas and Sim (2004)) that there is at least Γ such constraints. So, in that sense, the system becomes *more constrained* when Γ increases.

Hence, two opposite effects are at play: the number of constraints increases with Γ, intuitively making it more difficult to satisfy all of them, but the right-hand side they all have in common decreases, making it easier to satisfy all of them. This is why, in numerical examples, we observe a degree of diversification that first increases with Γ (the increase in the number of constraints is the dominant factor, so the increased diversification reflects the impossibility of satisfying all of the constraints by investing in fewer assets), before seeing it decrease steadily (the decrease in $\eta^*(\Gamma)$ then dominates the other trend, and it becomes possible to satisfy the constraints by investing in fewer assets.)

We also observe in experiments that, in the first phase where the degree of diversification increases with Γ, the ξ_i remain zero, as in the case with independent assets. In the second phase, however, when diversification decreases, we note that some of the ξ_i become non-zero. This situation was never observed when assets were independent; maximum diversification was then immediately followed by a complete drop in the number of assets invested in and the ξ_i were never zero as long as $\eta > 0$. In this special correlated case, when η becomes zero, $\xi_i = \sqrt{T}\, c \left(\sum_{j=1}^{n} Q_{ij}^{1/2}\, \chi_j\right)$ for all i and the robust problem becomes:

$$
\begin{aligned}
\max_{\chi} \quad & \sum_{i=1}^{n} \chi_i \ln k_i - \sum_{i=1}^{n} \sqrt{T}\, c \left(\sum_{j=1}^{n} Q_{ij}^{1/2}\, \chi_j\right) \\
\text{s.t.} \quad & \sum_{i=1}^{n} \chi_i = 1, \\
& \chi_i \geq 0, \ \forall i.
\end{aligned}
$$

As in the case of independent assets, the optimal solution of this linear programming problem is achieved at a corner point of the feasible set, i.e., when the budget of uncertainty is large enough so that $\eta = 0$, it becomes optimal to invest only in the asset with the highest worst-case return, here $\ln k_i - \sqrt{T}\, c \sum_{j=1}^{n} Q_{ji}^{1/2}$.

3.3 General Correlated Case

We now address the general case in the presence of correlation, when the coefficients of the square root of the correlation matrix can be positive or negative.

Theorem 3.5 (Optimal wealth and allocation)

(i) The optimal wealth in the robust portfolio management problem (9) is: $w_0 \exp(F(\Gamma))$, *where* F *is the function defined by:*

$$\begin{aligned} F(\Gamma) = \max_{\eta,\chi,\xi} \quad & \sum_{i=1}^{n} \chi_i \ln k_i - \eta\,\Gamma - \sum_{i=1}^{n} \xi_i \\ s.t. \quad & \eta + \xi_i - \sqrt{T}\, c \left(\sum_{j=1}^{n} Q_{ij}^{1/2}\, \chi_j \right) \geq 0, \;\; \forall i, \\ & \eta + \xi_i + \sqrt{T}\, c \left(\sum_{j=1}^{n} Q_{ij}^{1/2}\, \chi_j \right) \geq 0, \;\; \forall i, \\ & \sum_{i=1}^{n} \chi_i = 1, \\ & \eta \geq 0, \; \chi_i, \, \xi_i \geq 0, \;\; \forall i. \end{aligned} \tag{13}$$

(ii) The optimal amount of money invested at time 0 in stock i *is* $\chi_i\, w_0$, *for all* i, *where the* χ_i *are found by solving the linear programming problem (13).*

We see easily that the difference between Problem (12) and Problem (13) is that the constraints $\eta + \xi_i - \sqrt{T}\, c \left(\sum_{j=1}^{n} Q_{ij}^{1/2}\, \chi_j \right) \geq 0$ have become $\eta + \xi_i - \sqrt{T}\, c \left| \sum_{j=1}^{n} Q_{ij}^{1/2}\, \chi_j \right| \geq 0$ for all i.

We now provide the counterpart to Theorem 3.4 in the general case, which we state without proof.

Theorem 3.6 (Structure of the optimal solution) *Let* $S(\Gamma)$ *be the set of indices* j *such that* $\chi_j > 0$ *at optimality. There exists a set* $S_\xi^*(\Gamma)$ *of size* $|S(\Gamma)| - 1$ *such that* $(\chi_j)_{j \in S(\Gamma)}$ *is determined* *by the set of* $|S(\Gamma)|$ *equations:*

$$\begin{aligned} \left|\sum_{j\in S(\Gamma)} Q_{ij}^{1/2}\chi_j\right| &= \frac{\eta^*(\Gamma)}{\sqrt{T}c}, \ \forall i \in \mathcal{S}_\xi^*(\Gamma), \\ \sum_{j\in S(\Gamma)} \chi_j &= 1. \end{aligned}$$

For $j \notin S(\Gamma)$, $\chi_j = 0$.

Again, the robust optimization approach for the correlated case can be interpreted as choosing the sets $S(\Gamma)$ and $\mathcal{S}_\xi^*(\Gamma)$ appropriately.

Diversification. We have similar insights as in the special correlated case. In particular, the number of assets invested in first increases in Γ (we observe in numerical experiments that this corresponds to the case where $\eta > 0$ and $\xi_i = 0$ for all i), then decreases (this corresponds to the case where $\eta > 0$ but $\xi_i > 0$ for some i.) The main difference is that, when the decision-maker becomes very conservative (Γ is large), the number of assets invested in can stabilize at a number greater than 1. We can explain this behavior mathematically as follows. When $\eta = 0$, we combine the first two groups of constraints to obtain: $\xi_i = \sqrt{T}\,c\left|\sum_{j=1}^n Q_{ij}^{1/2}\,\chi_j\right|$ for all i. The robust problem is then:

$$\begin{aligned} \max_{\chi} \quad & \sum_{i=1}^n \chi_i \ln k_i - \sqrt{T}\,c \sum_{i=1}^n \left|\sum_{j=1}^n Q_{ij}^{1/2}\,\chi_j\right| \\ \text{s.t.} \quad & \sum_{i=1}^n \chi_i = 1, \\ & \eta \geq 0, \ \chi_i, \ \xi_i \geq 0, \ \forall i. \end{aligned}$$

This is a piecewise linear (not linear) programming problem, so the optimal solution will not necessarily be at a corner point of the feasible set, where the manager invested in only one asset.

4 Numerical Experiments

The purpose of this section is to compare the proposed Log-robust approach with the robust optimization approach that has been traditionally implemented in portfolio management. We will see that:

1. The Log-robust approach yields far greater diversification in the optimal asset allocation.

2. It outperforms the traditional robust approach, when performance is measured by percentile values of final portfolio wealth, if at least one of the following two conditions is satisfied: (a) the budget of uncertainty parameter is relatively small, or (b) the percentile considered is low enough.

This means that the Log-robust approach shifts the left tail of the wealth distribution to the right, compared to the traditional robust approach; how much of the whole distribution ends up being shifted depends on the choice of the budget of uncertainty. The rule of thumb in Bertsimas and Sim (2004), suggesting that the budget be of the order of the square root of the number of uncertain parameters, satisfies (a) and yields high-quality results in our experiments.

Setup.

The traditional robust approach when the stock prices belong to polyhedral uncertainty sets is due to Bertsimas and Sim (2004). The presence of correlation in real-life data requires extending their formulation to incorporate this case; the mathematical details are straightforward and left to the reader. The traditional framework, using the notations introduced at the beginning of the paper, is:

$$
\begin{aligned}
\max_{x,p,q,r} \quad & \sum_{i=1}^{n} x_i \exp\left[\left(\mu_i - \frac{\sigma_i^2}{2}\right) T\right] E\left[\exp\left(\sum_{j=1}^{n} Q_{ij}^{1/2} Z_j\right)\right] - \Gamma\, p - \sum_{i=1}^{n} q_i \\
\text{s.t.} \quad & \sum_{i=1}^{n} x_i = w_0, \\
& p + q_i \geq c\, r_i, \ \forall i, \\
& -r_i \leq \sum_{k=1}^{n} M_{ki}^{1/2} x_k \leq r_i, \ \forall i, \\
& p,\, q_i,\, r_i,\, x_i \geq 0, \ \forall i,
\end{aligned}
$$

with $M^{1/2}$ the square root of the covariance matrix of $\exp\left[\left(\mu_i - \frac{\sigma_i^2}{2}\right) T + \sqrt{T}\left(\sum_{j=1}^{n} Q_{ij}^{1/2} Z_j\right)\right]$.

In both the traditional robust and the Log-robust models, we downloaded six months' worth of daily stock price data for 50 stocks from Yahoo! Finance, computed the drift parameters and covariance matrix Q based on the continuously compounded rates of return $\ln(S_t/S_{t-1})$ and generated 1,000 scenarios for the stock prices six months from now. In the traditional model, we then used these 1,000 scenarios to compute $M^{1/2}$. Initial wealth is \$100,000. We take $c = 1.96$, which corresponds to a 95% confidence interval for a standardized Gaussian random variable. This choice (of a confidence interval with high probability rather than the full demand support) is common practice in the robust optimization literature to avoid over-conservatism. In particular, researchers routinely pick the confidence intervals as deviating from the nominal value by two standard deviations on each side (left and right). See for instance the references in Bertsimas and Thiele (2006) for numerical experiments reflecting these choices, performed by independent research teams. It appears that such an approach strikes a good trade-off between covering a large number of the potential values taken by the random variables, while not being too conservative. Ben-Tal et. al. (2006) explores in depth the consequences of this modeling, in particular with respect to having realizations fall outside the uncertainty set.

Analysis of optimal solution.

Figure 1 studies the level of diversification achieved in both models by showing the number of stocks invested in as a function of Γ. A key observation we make is that, while the numerical example in Bertsimas and Sim (2004) with artificial data (taken from Ben-Tal and Nemirovski (1999)) suggested that the robust approach would lead to a diversified portfolio for a wide range of budgets Γ, the results in Bertsimas and Sim (2004) appear to have been driven by the specific numerical values for the range forecasts of the stock returns (with tiny changes in mean and standard deviation from one stock to the next) and are not replicated with the real-life data we have considered. In our example, the traditional robust portfolio uses at most two stocks (and in general only one, as in the deterministic case). In contrast, the Log-robust model provides the manager with a diversified portfolio, with the number of stocks invested in increasing from 1 in the deterministic case ($\Gamma = 0$) to 31 for Γ between 21 and 25, and then decreasing steadily to 7 when $\Gamma = 43$. The number of stocks invested in then remains constant for higher values of Γ (the bump at $\Gamma = 45$ corresponds to a situation where two assets are virtually indistinguishable from a drift and variance perspective), up until the most conservative value of $\Gamma = 50$.

Figure 2 shows the number of shares bought in each of the stocks invested in either for $\Gamma = 10$ or $\Gamma = 20$, ranked in decreasing number of shares for $\Gamma = 10$. (Stocks that are not invested in in either case are not shown.) In the deterministic model and the Traditional-Robust model, the manager only invests in Air Products and Chemicals, Inc. (APD). Figure 2 indicates in particular

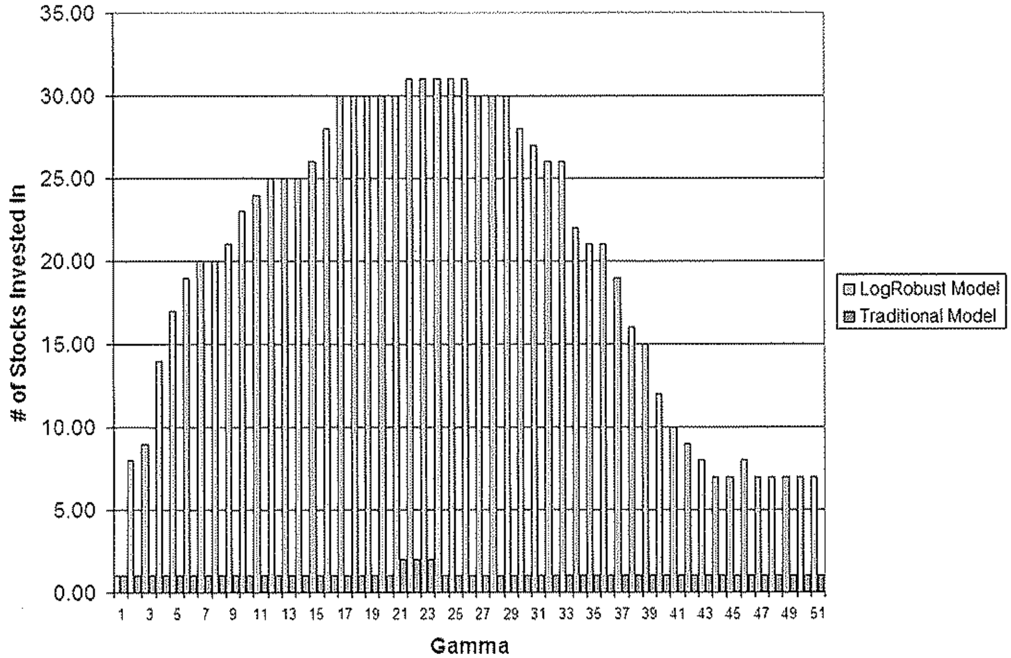

Figure 1: Number of stocks in optimal portfolio for Γ varying from 0 to 50, in the Traditional-Robust and Log-robust models.

that the number of shares bought in each stock is often quite substantial, so the diversification effect observed in Figure 1 is not due to the manager buying just one or two shares of more stocks; in other words, we achieve genuine diversification.

Analysis of performance in simulations.

Since the goal of the proposed methodology is to protect against downside risk, we pay particular attention to the 99% and 95% Value-at-Risk of the portfolio in the Traditional-Robust and the Log-Robust models. We gather data on the other percentiles as well, to study under which circumstances the framework proposed here outperforms the traditional one. The simulations were performed using @Risk 4.5 from Palisade Corporation. We consider two cases: (i) the case where the random variables do obey a Normal distribution, and the only mistake made by the manager implementing the traditional robust approach is that he uses symmetric confidence intervals for the stock prices rather than for the true drivers of uncertainty, their continuously compounded rates of change, (ii) the case where the random variables have "fat tails", as has been

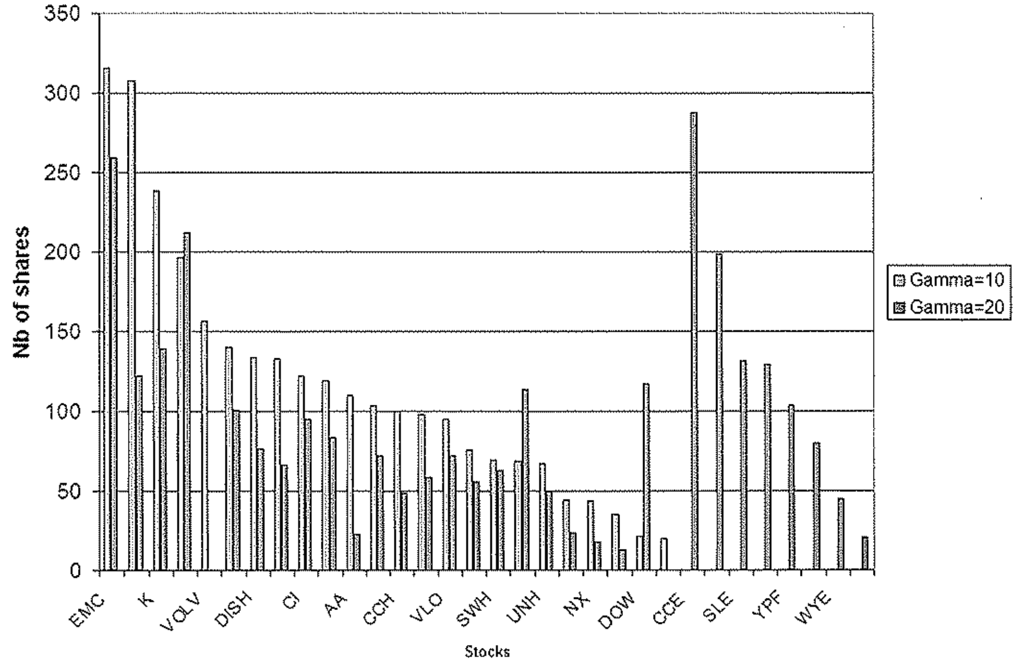

Figure 2: Number of shares in optimal portfolio for $\Gamma = 10$ and $\Gamma = 20$, for various stocks, in the Log-robust model.

observed in practice (see Hull 2002), in which case the Lognormal model of stock price behavior underestimates rare events. This happens for instance when the scaled random variables Z obey a Logistic distribution. To calibrate the distribution we selected the same 95% confidence interval as that given by the Gaussian model with mean 0 and standard deviation 1, to keep the same range forecasts throughout. (The methodology was tested for other distributions as well and yielded similar results.)

Recall that the 99% and 95% Value-at-Risk are the 1% and the 5% percentiles of the portfolio wealth. For instance, the 99% VaR is the number such that there is only a 0.01 probability that the portfolio value will fall below that number. The decision-maker naturally wants these worst-case portfolio values to be as large as possible, so that investors remain wealthy even under adverse market conditions. In particular, 99% and 95% VaR are *risk-adjusted performance measures*, not risk measures, and should be maximized, not minimized.

Normal distribution.

Table 1 keeps track of the 99% Value-at-Risk for values of Γ varying from 0 to 50 in increments of 5, in the traditional and the Log-robust models; the last column shows the relative gain in 99% VaR when the manager implements the Log-robust approach. The values are obtained using 10,000 replications. We observe that 99% VaR decreases steadily as the level of conservatism

Γ	Traditional	Log-Robust	Relative Gain
5	70958.81	107828.94	51.96%
10	70958.81	104829.93	47.73%
15	70958.81	102502.79	44.45%
20	70958.81	101707.00	43.33%
25	70958.81	100905.96	42.40%
30	70958.81	101763.58	43.41%
35	70958.81	98445.23	38.74%
40	70958.81	96120.18	35.46%
45	70958.81	94253.62	32.83%
50	70958.81	94032.09	32.52%

Table 1: 99% VaR as a function of Γ for Gaussian distribution.

(measured by the budget Γ) increases, so that the relative gain from using the Log-robust approach decreases from about 52% to about 33%. (Recall that for $\Gamma = 0$, both frameworks yield the deterministic model.) Because the stocks are correlated, the uncertain parameter z_i affects not only the stock price of asset i, but also the stock prices of the other assets. This is why, although the decision-maker invests in at most 31 stocks, the VaR keeps decreasing – instead of becoming constant – for values of Γ greater than 31. Bertsimas and Sim (2004) have suggested selecting a value for the budget of uncertainty of the order of $\sqrt{n}$ (about 7 here) with n the number of uncertainty drivers, and Table 1 suggests that values of Γ in the 5-10 range are precisely those that maximize the benefit of using the Log-robust approach, at least for the 99% VaR. We investigate this point further in Figure 3, which shows the relative gain of the Log-robust approach for percentiles of the portfolio value between 5 and 95%, in increments of 5, and Γ between 0 and 50, in increments of 5. (Negative relative gains indicated that the traditional robust approach performs better.)

We observe that the relative performance of the Log-robust approach decreases as Γ increases and as the percentile increases; up to the 10th percentile (90% VaR), the Log-Robust model outperforms the traditional approach for any value of Γ. The Log-robust approach performs best for a risk-averse decision-maker (focusing on 99% or 95% VaR) and for moderate values of Γ

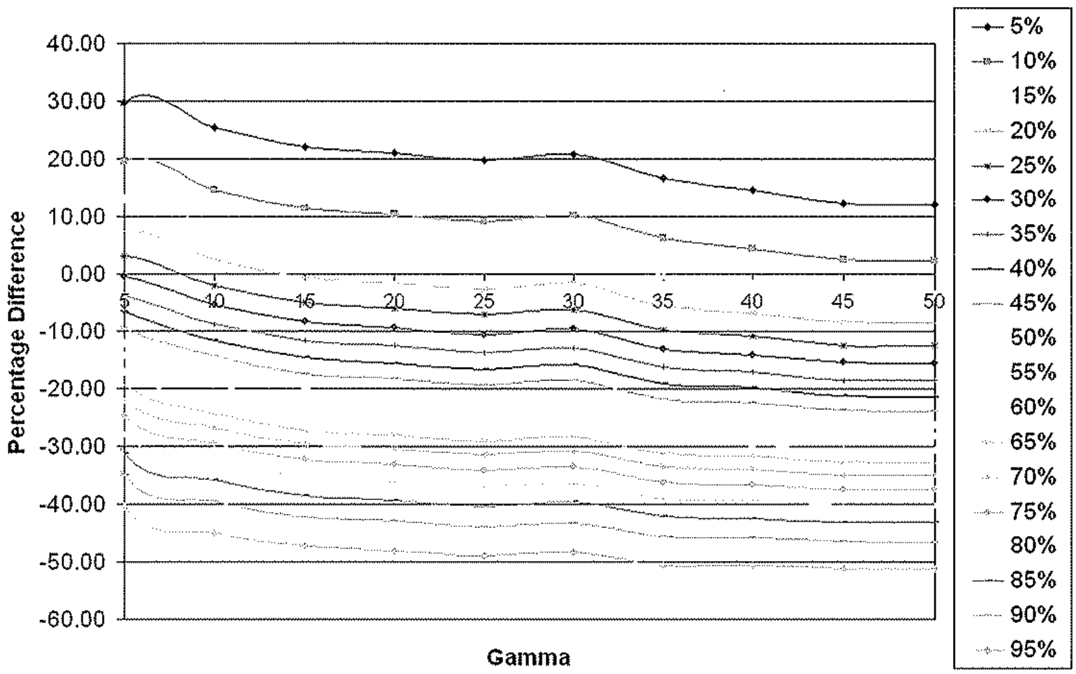

Figure 3: Relative gain of the Log-robust model compared to the Traditional robust model, for percentiles from 5% to 95% and Γ from 0 to 50, in the Gaussian case.

(about 5 or 10).

Logistic distribution.

We now consider the more realistic case where the distribution has "fat tails," i.e., the Gaussian assumption underlying the Lognormal model underestimates the risk of extreme events. Table 2 keeps track of the 99% Value-at-Risk for values of Γ varying from 0 to 50 in increments of 5, in the traditional and the Log-robust models, when the scaled random variables Z obey Logistic distributions with 95% confidence intervals $[-c, c]$; as in Table 1, the last column shows the relative gain in 99% VaR when the manager implements the Log-robust approach. As in the Gaussian case, the relative performance of the Log-robust approach decreases as Γ increases and as the percentile increases; up to the 10th percentile (90% VaR), the Log-Robust model outperforms the traditional approach for any value of Γ. The Log-robust approach performs best for a risk-averse decision-maker (focusing on 99% or 95% VaR) and for moderate values of Γ

Γ	Traditional	Log-Robust	Relative Gain
5	68415.97	108234.32	58.20%
10	68415.97	105146.66	53.69%
15	68415.97	102961.66	50.49%
20	68415.97	102124.75	49.27%
25	68415.97	101294.347	48.06%
30	68415.97	102206.73	49.39%
35	68415.97	98508.69	43.98%
40	68415.97	95940.01	40.23%
45	68415.97	93841.05	37.16%
50	68415.97	93562.59	36.76%

Table 2: 99% VaR as a function of Γ for Logistic distribution.

(about 5 or 10). This is shown in Figure 4. The changes compared to Figure 3 are minor; for instance, the relative gain for $\Gamma = 5$, considering the 20th percentile, has changed from 7.41% (Gaussian case) to 6.05% (Logistic case).

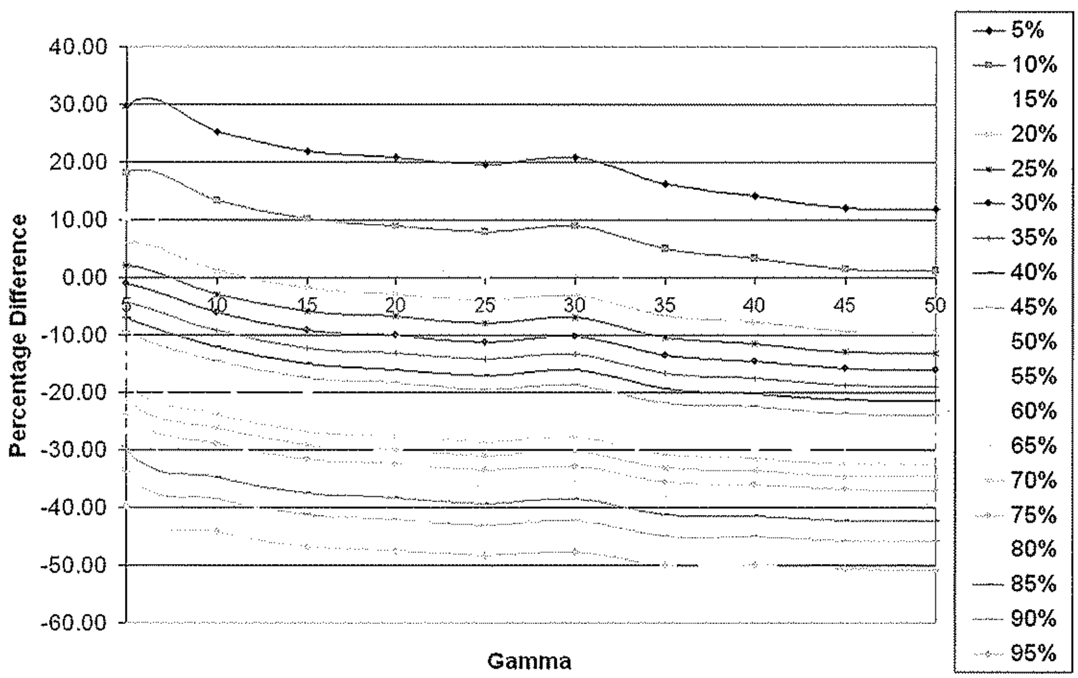

Figure 4: Relative gain of the Log-robust model compared to the Traditional robust model, for percentiles from 5% to 95% and Γ from 0 to 50, in the Logistic case.

Conclusions of Experiments.

Our numerical results indicate that incorporating robustness at the level of the true uncertainty driver, the continuously compounded rate of return, results in better performance for the risk-averse manager maximizing his 99% VaR (or 95% or 90% VaR). They also suggest that the budget of uncertainty should be of the order of the square root of the random variables to optimize the performance of the approach. This is in line with rules of thumb available in the literature.

5 Conclusions

We have proposed a robust optimization approach to portfolio management, where robustness is incorporated in the continuously compounded rates of return of the stock prices rather than in the prices themselves. This departure from the traditional robust framework aligns our model with the finance literature without requiring the mathematically convenient assumption of stock prices following a Lognormal process, which has been shown to underestimate extreme events in practice. We have obtained a robust formulation that is linear and thus can be solved efficiently, and have derived theoretical insights into the worst-case uncertainty and the optimal number of shares to buy of each stock. In numerical experiments when the decision-maker maximizes his 95% or 99% Value-at-Risk, the Log-robust approach outperforms the traditional robust optimization approach by double-digit margins, with an even more significant gain if the budget of uncertainty is well-chosen (about the square root of the number of stocks). This is because the traditional robust optimization approach does not achieve diversification for real-life financial data. Hence, we believe the Log-robust approach holds much potential in portfolio management under uncertainty.

Acknowledgments

We would like to thank two anonymous reviewers for their comments, which have helped us clarify and strengthen the paper.

References

Al Najjab, M., A. Thiele. (2007) The Log-Logistic option pricing model. Technical Report, Lehigh University, Department of Industrial and Systems Engineering, Bethlehem, PA.

Ben-Tal, A., A. Nemirovski. (1999) Robust solutions to uncertain programs. *Oper. Res. Letters*, 25, 1-13.

Ben-Tal, A., S. Boyd, A. Nemirovski. (2006) Extending Scope of Robust Optimization: Comprehensive Robust Counterparts of Uncertain Problems. *Math. Prog. B,* 107, 63-89.

Bertsekas, D. (1999) *Nonlinear Programming* 2nd ed. Athena Scientific. Belmont, MA.

Bertsimas, D., D. Pachamanova. (2008) Robust Multiperiod Portfolio Management with Transaction Costs. *Computers and Operations Research*, 35(1), 3-17.

Bertsimas, D., M. Sim. (2004) The Price of Robustness. *Operations Research*, 52(1): 35-53.

Bertsimas, D., A. Thiele. (2006) Robust and data-driven optimization: modern decision-making under uncertainty, *INFORMS Tutorials in Operations Research,* Chapter 4.

Black, F., M. Scholes. (1973) The pricing of options and corporate liabilities. *Journal of Political Economy,* 81: 637-59.

Blattberg, R., N. Gonedes. (1974) A Comparison of the Stable and Student Distributions as Statistical Models for Stock Prices. *Journal of Business,* 47, 244-80.

Boyd, S., L. Vandenberghe. (2004) *Convex Optimization* Cambridge University Press. Cambridge, UK.

Chopra, V., W. Ziemba. (1993) The effect of errors in means, variances, and covariances on optimal portfolio choice. *Journal of Portfolio Management*, 19, 6-11.

Cont, R. (2001) Empirical Properties of Asset Returns – Stylized Facts and Statistical Issues. *Quantitative Finance,* 223-236.

Fabozzi, F., P. Kolm, D. Pachamanova, S. Focardi. (2007) *Robust Portfolio Optimization and Management.* Wiley.

Fama, E. (1965) The Behavior of Stock Prices. *Journal of Business,* 38, 34-105.

Erdogan, E., D. Goldfarb, G. Iyengar. (2004) Robust Portfolio Management. Technical Report, Department of Industrial Engineering and Operations Research, Columbia University, New York, NY.

Goldfarb, D., G. Iyengar. (2003) Robust Portfolio Selection Problems. *Mathematics of Operations Research,* 28, 1-38.

Hull, J. C. (2002) *Options, Futures and Other Derivatives.* 5th edition. Prentice Hall. Upper Saddle River, NJ.

Jansen, D., C. deVries. (1991) On the Frequency of Large Stock Returns – Putting Booms and Busts into Perspective. *The Review of Economics and Statistics,* 73(1), 18-24.

Kon, S. (1984) Models of Stock Returns – A Comparison. *Journal of Finance,* 39, 147-65.

Markowitz, H. M. (1959) *Portfolio selection: efficient diversification of investments.* Wiley, New York.

Pachamanova, D. (2006) Handling Parameter Uncertainty in Portfolio Risk Minimization. *Journal of Portfolio Management,* 32(4), 70-78.

Richardson, M., T. Smith. (1993) A Test for Multivariate Normality in Stock Returns. *Journal of Business,* 66, 295-321.

A Proofs

A.1 Proof of Lemma 2.1

Because we do not allow short sales, the coefficient in front of the exponential is nonnegative and the exponential is minimized for the smallest value of its argument. □

A.2 Proof of Lemma 2.2

(i) Problem (4) is an unconstrained convex optimization problem, and as such its optimal solution is found by setting the gradient of the objective to zero. (ii) follows from complementary slackness applied to Problem (3). (iii) If $x_i = 0$ and $z_i > 0$, then taking $z_i' = 0$ yields a feasible solution with same objective value. □

A.3 Proof of Theorem 2.3

Injecting Lemma 2.2 into Equation (4), and using strong duality in convex programming with Slater's condition (see Bertsekas 1999), we obtain that the robust portfolio management problem (1) is equivalent to:

$$\begin{aligned}
\max_{x,\,\alpha,\,\lambda^0,\,\lambda^1} \quad & \sum_{i=1}^{n} \left(\frac{\alpha - \lambda_i^0 + \lambda_i^1}{\sigma_i \sqrt{T}\, c} \right) \cdot \left[1 + \ln \left(\frac{x_i\, k_i\, \sigma_i \sqrt{T}\, c}{\alpha - \lambda_i^0 + \lambda_i^1} \right) \right] - \alpha\, \Gamma - \sum_{i=1}^{n} \lambda_i^1 \\
\text{s.t.} \quad & \sum_{i=1}^{n} x_i = w_0, \\
& \alpha \geq 0,\ \lambda_i^0,\ \lambda_i^1,\ x_i \geq 0,\ \forall i,
\end{aligned}$$

or alternatively, using the change of variable: $\beta_i = \dfrac{\alpha - \lambda_i^0 + \lambda_i^1}{\sigma_i \sqrt{T}\, c}$, which must be non-negative due to the term in log:

$$\begin{aligned}
\max_{x,\,\alpha,\,\beta,\,\lambda^1} \quad & \sum_{i=1}^{n} \beta_i \cdot \left[1 + \ln \left(\frac{x_i\, k_i}{\beta_i} \right) \right] - \alpha\, \Gamma - \sum_{i=1}^{n} \lambda_i^1 \\
\text{s.t.} \quad & \sum_{i=1}^{n} x_i = w_0, \\
& \alpha + \lambda_i^1 - \sigma_i \sqrt{T}\, c\, \beta_i \geq 0,\ \forall i, \\
& \alpha \geq 0,\ \beta_i,\ \lambda_i^1,\ x_i \geq 0,\ \forall i.
\end{aligned} \tag{14}$$

We solve Problem (14) by first maximizing over the x_i and then over the remaining variables. The maximizing problem in the x_i can be formulated as:

$$\begin{aligned}
\max_{x} \quad & \sum_{i=1}^{n} \beta_i \cdot \ln x_i \\
\text{s.t.} \quad & \sum_{i=1}^{n} x_i = w_0, \\
& x_i \geq 0,\ \forall i.
\end{aligned} \tag{15}$$

Problem (15) is a convex optimization problem, which we solve using a Lagrange approach, obtaining $x_i = \frac{\beta_i\, w_0}{\sum_{j=1}^{n} \beta_j}$. (Note that this means that x_i and β_i are both zero or both positive, for each i.) We reinject the optimal asset allocation into Problem (14) and now have to solve:

$$\begin{aligned} \max_{\alpha,\beta,\lambda^1} \quad & \sum_{i=1}^{n} \beta_i \cdot \left[1 + \ln\left(\frac{w_0\, k_i}{\sum_{j=1}^{n} \beta_j}\right)\right] - \alpha\,\Gamma - \sum_{i=1}^{n} \lambda_i^1 \\ \text{s.t.} \quad & \alpha + \lambda_i^1 - \sigma_i \sqrt{T}\, c\, \beta_i \geq 0, \ \forall i, \\ & \alpha \geq 0,\ \beta_i,\ \lambda_i^1 \geq 0,\ \forall i. \end{aligned} \tag{16}$$

Because the right-hand side of the feasible set of Problem (16) is zero, we can parametrize over $\theta \geq 0$ where $\sum_{j=1}^{n} \beta_j = \theta$ (note that θ must be nonnegative for the logarithm to be defined) and scale the decision variables by $1/\theta$. Problem (16) becomes:

$$\begin{aligned} \max_{\theta}\ \theta \cdot \max_{\eta,\chi,\xi} \quad & \sum_{i=1}^{n} \chi_i \cdot \left[1 + \ln\left(\frac{w_0\, k_i}{\theta}\right)\right] - \eta\,\Gamma - \sum_{i=1}^{n} \xi_i \\ \text{s.t.} \quad & \eta + \xi_i - \sigma_i \sqrt{T}\, c\, \chi_i \geq 0,\ \forall i, \\ & \sum_{i=1}^{n} \chi_i = 1, \\ & \eta \geq 0,\ \chi_i,\ \xi_i \geq 0,\ \forall i. \end{aligned}$$

(Note that, with these new notations, $x_i = \chi_i w_0$ for all i.) We then regroup the terms depending on θ and use $\sum_{i=1}^{n} \chi_i = 1$ to reformulate the robust optimization problem as:

$$\max_{\theta \geq 0} \left[1 + F(\Gamma) + \ln\left(\frac{w_0}{\theta}\right)\right] \theta, \tag{17}$$

where F is defined by Equation (7). The objective in Problem (17) is concave, as is easily checked by computing the second derivative, and the optimal value of θ follows by setting the first derivative to zero. This yields: $\theta = w_0 \exp(F(\Gamma))$. Reinjecting into the objective leads to an optimal wealth of $w_0 \exp(F(\Gamma))$. $\square$

A.4 Proof of Corollary 2.4

We use that $z_i = (\ln k_i - F(\Gamma))/(\sigma_i \sqrt{T}\, c)$ as stated in the remarks after Theorem 2.3; this equation was derived by injecting into Equation (5) the fact that $x_i = w_0\, \chi_i = w_0\, \beta_i/\theta$ at optimality, with $\theta = w_0 \exp(F(\Gamma))$ and $\beta_i = (\alpha - \lambda_i^0 + \lambda_i^1)/(\sigma_i \sqrt{T}\, c)$. Since it is strictly suboptimal to invest in only one stock by assumption, we have that:

$$F(\Gamma) > \ln k_i - \sigma_i \sqrt{T}\, c, \ \forall i,$$

obtained by computing the objective value of Problem (7) for the feasible solution $\chi_i = 1$, $\chi_j = 0$ for $j \neq i$, $\eta = 0$, $\xi_i = \sigma_i \sqrt{T}\, c$ and $\xi_j = 0$ for $j \neq i$.

Now, consider the dual of Problem (7):

$$\begin{aligned} \min_{u,v} \quad & v \\ \text{s.t.} \quad & \sigma_i \sqrt{T}\, c\, u_i + v \geq \ln k_i, \ \forall i, \\ & \sum_{i=1}^{n} u_i \leq \Gamma, \\ & 0 \leq u_i \leq 1, \ \forall i. \end{aligned} \tag{18}$$

By complementarity slackness, we know that for all i such that $x_i > 0$, i.e., $\chi_i > 0$, we have $\sigma_i \sqrt{T}\, c\, u_i + v = \ln k_i$ so that $u_i = (\ln k_i - v)/(\sigma_i \sqrt{T}\, c)$ for all i such that $\chi_i > 0$.

But it is easy to see that: $\sum_{i \in S(\Gamma)} u_i = \Gamma$ at optimality, where $S(\Gamma)$ is the set of assets the manager invests in. This is because, if $\sum_{i=1}^{n} u_i \leq \Gamma$ is not tight at optimality, we can solve Problem (18) after discarding this constraint, which yields a strategy of investing everything in that asset i achieving the smallest $\ln k_i - \sigma_i \sqrt{T}\, c$. This contradicts the assumption that it is strictly suboptimal to invest in only one stock.

Note that $v = F(\Gamma)$ by strong duality. Using that $\sum_{i \in S(\Gamma)} u_i = \Gamma$ and rearranging leads to:

$$F(\Gamma) = \frac{\sum_{i \in S(\Gamma)} \frac{\ln k_i}{\sigma_i \sqrt{T}\, c} - \Gamma}{\sum_{i \in S(\Gamma)} \frac{1}{\sigma_i \sqrt{T}\, c}}.$$

So, to have $z_j = 0$ for some j such that $x_j > 0$, we must have $F(\Gamma) = \ln k_j$ or, using that $\ln k_j = (\mu_j - \sigma_j^2/2)T$:

$$\Gamma = \sum_{i \in S(\Gamma)} \frac{\left[(\mu_i - \frac{\sigma_i^2}{2}) - (\mu_j - \frac{\sigma_j^2}{2}) \right] T}{\sigma_i \sqrt{T}\, c}.$$

But the right-hand side does not depend on Γ (except in the choice of indices), so the equality cannot possibly hold.

It follows immediately that $0 < z_i < 1$ for all i such that $x_i > 0$. $\square$

A.5 Proof of Theorem 2.5

The result is trivial if we only invest in one asset. If $S(\Gamma)$ has at least two elements, we use Corollary 2.4 to justify that $0 < z_i < 1$ for all i such that $x_i > 0$. Let $R = \sum_{i \in S(\Gamma)} \frac{1}{\sigma_i \sqrt{T} c}$. The robust portfolio management problem is equivalent to the convex optimization problem:

$$
\begin{aligned}
\max_{x} \quad & R \cdot \exp\left(-\frac{\Gamma}{R}\right) \cdot \left[\prod_{i \in S(\Gamma)} (x_i k_i \sigma_i \sqrt{T} c)^{\frac{1}{\sigma_i \sqrt{T} c}} \right]^{\frac{1}{R}} \\
\text{s.t.} \quad & \sum_{i \in S(\Gamma)} x_i = w_0.
\end{aligned}
\tag{19}
$$

The proof is as follows. We use the first bullet point of Lemma 2.2 (ii) to rewrite the objective function of Problem (3) as $\sum_{i \in S(\Gamma)} \frac{\alpha}{\sigma_i \sqrt{T} c} = \alpha\, R$ where R has been defined above. To find α we use that $\Gamma = \sum_{i \in S(\Gamma)} z_i^*$:

$$
\begin{aligned}
\Gamma &= \sum_{i \in S(\Gamma)} \frac{1}{\sigma_i \sqrt{T} c} \left[\ln \left(\frac{x_i\, k_i\, \sigma_i \sqrt{T} c}{\alpha} \right) \right] \\
&= \sum_{i \in S(\Gamma)} \frac{1}{\sigma_i \sqrt{T} c} \ln(x_i\, k_i\, \sigma_i \sqrt{T} c) - \ln \alpha \cdot \sum_{i \in S(\Gamma)} \frac{1}{\sigma_i \sqrt{T} c}.
\end{aligned}
$$

This yields:

$$
\alpha = \exp\left(\frac{-\Gamma}{R}\right) \cdot \left[\prod_{i \in S(\Gamma)} (x_i\, k_i \sigma_i \sqrt{T} c)^{\frac{1}{\sigma_i \sqrt{T} c}} \right]^{\frac{1}{R}}. \tag{20}
$$

We then inject Equation (20) into the objective $\alpha\, R$ to obtain Problem (19), which is convex because the geometric mean is a concave function of its arguments.

We now prove Theorem 2.5. To find the optimal value of the x_i's in Problem (19), we invoke the convexity of the problem and introduce the Lagrangian multiplier δ of the budget constraint; the model becomes:

$$
\max_{x} \ \alpha(x) R - \delta \left(\sum_{i \in S(\Gamma)} x_i - w_0 \right).
$$

We set the gradient of the objective to zero to find the optimum value of the x_i's. This yields, for all i in $S(\Gamma)$,

$$\exp\left(-\frac{\Gamma}{R}\right)\cdot\left(\prod_{j\in S(\Gamma)}(x_j k_j \sigma_j \sqrt{T}c)^{\frac{1}{\sigma_j\sqrt{T}c}}\right)^{\frac{1}{R}} = \delta\,(x_i\,\sigma_i\,\sqrt{T}c). \tag{21}$$

So $x_i\,\sigma_i$ is constant for all i such that $x_i > 0$.

We now prove that the set $S(\Gamma)$ is of the type $\{1,\ldots,j\}$ for some j, where the assets have been ranking in decreasing order of their nominal return.

Setting the variables η and ξ_i for all i to their optimal value $\eta^*(\Gamma)$ and $\xi_i^*(\Gamma)$ in Problem (7) yields the following linear programming problem in the χ_i:

$$\begin{aligned}
\max_{\chi}\quad & \sum_{i=1}^{n}\chi_i\,\ln k_i\\
\text{s.t.}\quad & \sum_{i=1}^{n}\chi_i = 1,\\
& 0\le \chi_i \le \frac{\eta^*(\Gamma)+\xi_i^*(\Gamma)}{\sigma_i\sqrt{T}c},\ \forall i.
\end{aligned}$$

Because this is a linear programming problem, it can be solved by dualizing only some of the constraints (here, the coupling constraint) and keeping the others (here, the bound constraints) in the feasible set. Let a be the dual variable associated with the coupling constraint. We obtain:

$$\begin{aligned}
\max_{\chi}\quad & \sum_{i=1}^{n}\chi_i\,\ln k_i - a\left(\sum_{i=1}^{n}\chi_i - 1\right)\\
\text{s.t.}\quad & 0\le \chi_i \le \frac{\eta^*(\Gamma)+\xi_i^*(\Gamma)}{\sigma_i\sqrt{T}c},\ \forall i.
\end{aligned}$$

At optimality, χ_i is equal to its lower bound if the coefficient in front of χ_i in the objective, $\ln k_i - a$, is negative, and to its upper bound if that coefficient is non-negative. □

A.6 Proof of Lemma 3.1

The feasible set is convex (all the constraints are less-than-or-equal to constraints with convex functions in the left-hand side) and the objective is the weighted sum with nonnegative coefficients of the composition of convex functions with affine functions of the decision variables (Boyd and Vandenberghe 2004). □

A.7 Proof of Lemma 3.2

Follows immediately from solving the Lagrange relaxation of Problem (11) and invoking strong duality in convex optimization, since Slater's condition is satisfied. □

A.8 Proof of Theorem 3.3

Similar to that of Theorem 2.3. □

A.9 Proof of Theorem 3.5

The proof is similar to that of Theorem 2.3 and we only sketch the main ideas. We use the transformation: $\tilde{y}_j = y_j^+ - y_j^-$ (and hence, $|\tilde{y}_j| = y_j^+ + y_j^-$), which does not change the optimal objective because Problem (10) is convex. This yields, for the inner minimization problem:

$$\begin{aligned} \min_{y^+, y^-} \quad & \sum_{i=1}^{n} x_i\, k_i \exp\left[\sqrt{T}c \sum_{j=1}^{n} Q_{ij}^{1/2} (y_j^+ - y_j^-)\right] \\ \text{s.t.} \quad & \sum_{j=1}^{n} (y_j^+ + y_j^-) \leq \Gamma, \\ & y_j^+ + y_j^- \leq 1, \ \forall j, \\ & y_j^-, y_j^+ \geq 0, \ \forall j. \end{aligned} \tag{22}$$

We solve the convex optimization problem (22) using a Lagrange approach, with Lagrange multipliers α, λ_i^1, λ_i^{-0} and λ_i^{+0} for all i. Setting the gradient to zero yields:

$$2(\alpha + \lambda_j^1) - \lambda_j^{+0} - \lambda_j^{-0} = 0, \ \forall j,$$

and:

$$y_j^+ - y_j^- = \frac{1}{\sqrt{T}c} \sum_{i=1}^{n} Q_{ij}^{-1/2} \ln\left(\frac{[\mathbf{Q}^{1/2}(\alpha \mathbf{e} + \boldsymbol{\lambda}^1 - \boldsymbol{\lambda}^{-0})]_i}{\sqrt{T}c\, x_i\, k_i}\right), \ \forall j.$$

Using that $\lambda_j^{+0} = 2(\alpha + \lambda_j^1) - \lambda_j^{-0}$, introducing the change of variables: $\sqrt{T}c\boldsymbol{\beta} = \mathbf{Q}^{-1/2}(\alpha \mathbf{e} + \boldsymbol{\lambda}^1 - \boldsymbol{\lambda}^{-0})$, injecting the nonnegativity of λ_j^{-0} and λ_j^{+0} for all j and scaling by $\sum_{j=1}^{n} \beta_j$ yields the desired result, similarly to the proof of Theorem 2.3. □