CARISMA

A multi-disciplinary research centre focussed on understanding, modelling, quantification, management and control of RISK

- School of Information Systems, Computing and Mathematics
- Brunel Business School
- School of Engineering
- School of Social Sciences

TECHNICAL REPORT

CTR/76-08

Robust Optimisation and Portfolio and Selection:
The Cost of Robustness

Christine Gregory, Gautam Mitra, Ken Darby-Dowman

CARISMA

Centre for the Analysis of Risk and Optimisation Modelling Applications

Brunel UNIVERSITY WEST LONDON

Robust Optimization and Portfolio Selection: The Cost of Robustness

Christine Gregory   Ken Darby-Dowman   Gautam Mitra

September 23, 2008

Abstract

Robust optimization is a tractable alternative to stochastic programming particularly suited for problems in which parameter values are unknown, variable, and their distributions are uncertain. We evaluate the cost of robustness of the robust counterpart to the maximum return portfolio optimization problem. The uncertainty of asset returns is modelled by polyhedral uncertainty sets as opposed to the earlier proposed ellipsoidal sets. We derive the robust model from a min-regret perspective and examine the properties of robust models with respect to portfolio composition. We investigate the effect of different definitions of the bounds on the uncertainty sets and show that robust models yield well diversified portfolios, in terms of the number of assets and asset weights.

Keywords: robust optimization; portfolio selection; mixed integer robust optimization

1 Introduction

Robust optimization is a min-regret modelling methodology that seeks to minimise the negative impact of future events when the values of model parameters are unknown, variable, and their distributions are uncertain. Variability, commonly expressed as a statistical measure (e.g. standard deviation, variance), is the "naturally occurring, unpredictable change" (Burgman, 2005) of a parameter and is not reducible by the acquisition of more knowledge (Vose, 2000). That is, no matter how much information is acquired about the parameter it is not going to change its behaviour. Uncertainty, however, reflects a lack of knowledge about a future event and can be reduced (but not necessarily eliminated) by gathering more information (Vose, 2000); for example, by collecting more data, parameter distributions can be more precisely estimated. To illustrate the difference between variability and uncertainty, consider the random walk of a stock price, which depicts its "naturally occurring, unpredictable change". It is not possible to change the random walk, i.e. make it less volatile, no matter how much data is collected or knowledge acquired regarding its past behaviour – this is variability. However, if we estimate the distribution of a random walk, acquiring more information about its past behaviour will increase the precision of that estimate – this is uncertainty. Furthermore, throughout this paper, a model in which the parameter values are unknown, variable, and their distributions uncertain, is referred to as an uncertain model.

Model robustness has several different definitions, depending on whether model constraints are hard constraints or can be violated with some penalty. In this paper, we are only concerned with the optimization of mathematical models with hard constraints. Therefore, we define a model as robust if the optimal objective is achieved or exceeded with high probability and the solution remains feasible for all possible realizations of the unknown parameter values within an uncertainty set, even if the assumed distributions and estimates of the parameters are imprecise. Although parameter values are unknown, historical data (if available) may be used to estimate the uncertainty set. It does not need to encapsulate every possible realization of the parameter, but only the "most likely" values, the specification of which is partially a subjective decision. The two most common ways of defining the geometry of uncertainty sets are polyhedral and ellipsoidal sets, which are discussed further in section 2.

There are three common approaches for dealing with uncertain mathematical programs: 1) post-optimal analysis, such as sensitivity analysis, 2) treating uncertainty in the modelling stage, such as stochastic programming and robust optimization, and 3) ignoring it and applying deterministic models. Sensitivity analysis asks the question: How sensitive is the optimal solution to changes in the parameters? In other words, how much can the actual parameter values differ from their estimates before the solution loses optimality? It is a means of studying the effects of variability and uncertainty on the optimal decision but does not protect against them. Alternatively, stochastic programming is a well-known methodology that treats variability and uncertainty in the modelling stage. Parameter distributions, or a model that accurately creates them, is assumed to be known. In this paradigm, constraints are allowed to be violated, but with a specified penalty. Therefore, the solution may not be feasible for all scenarios. The idea is to hedge against the risk of unfavourable scenarios that may occur in the future. A difficulty with stochastic programming is that as the number of scenarios increases, the computational demands increase significantly. In addition, the quality of the solution is determined by the validity of the assumptions governing the stochastic process used to generate scenarios. Finally, a deterministic approach that assigns static estimates to unknown parameter values may yield results that are unreliable and unusable. If the realized parameter values deviate too much from their estimates, constraints are violated and decisions become infeasible.

During the last ten years, robust optimization, a tractable alternative to stochastic programming, has been applied to portfolio selection and is of growing interest. The robust counterpart is a deterministic formulation of the stochastic problem, which optimizes an objective such that all constraints are satisfied for all possible values of each uncertain parameter defined within a set. Unlike stochastic programming, it does not rely on knowing the exact distributions of returns – which are rarely known in practice and typically estimated. Underpinning robust optimization is a desire for mathematical models producing solutions insensitive to changes in uncertain parameters such that a) it is computationally manageable, b) decisions are useable – if input data changes, the solution is near optimal with high probability, and c) the robustness of the solution is worth the sacrifice of optimality.

The purpose of this paper is to analyze the behaviour, robustness, and cost of the robust counterpart formulation of the portfolio optimization problem purposed by

Bertsimas and Sim (2004), in which unknown parameters were modelled by polyhedral sets. Nearly all previous robust portfolio optimization research has stemmed from the work of Ben-Tal and Nemirovski (1999) who model unknown parameters by ellipsoidal sets (discussed further in section 2). It has been suggested that polyhedral sets are a crude model of uncertainty and the resulting robust counterpart is too simplistic (Ben-Tal and Nemirovski, 1998). However, it is worthwhile investigating whether or not the robust portfolio optimization model in question produces quality solutions, even though it is a more simplistic formulation. To our knowledge, an empirical investigation of this model, its properties (i.e. portfolio composition), an analysis of the cost of robustness, and the effects of changing the size of the uncertainty set has not been carried out before.

In section 2, we discuss the motivation behind robust optimization. In section 3, we highlight the literature supporting robust portfolio optimization. Section 4 shows the derivation of the linear robust counterpart purposed by Bertsimas and Sim (2004) and an interpretation of the robust counterpart. In section 5, we discuss the Bertsimas and Sim model for portfolio optimization and the composition of robust portfolios. In section 6, we introduce measures of robustness and the cost of robustness and evaluate robust portfolios based on these measures. Lastly, in section 7, we discuss the practical applicability of this research and suggest further investigations.

2 Robust Decisions of Uncertain Mathematical Programs

An uncertain LP can be expressed as follows, (Ben-Tal and Nemirovski, 1998):

$$\begin{aligned} \text{Max} \quad & c^T x \,. \\ \text{S.t.} \quad & Ax \le b \,, \end{aligned} \tag{1}$$

where A and b are the unknown and variable parameters belonging to an uncertainty set U. The robust counterpart is given as

$$\begin{aligned} \text{Max} \quad & c^T x \,. \\ \text{S.t.} \quad & Ax \le b \,, \qquad \forall (A,b) \in U \,. \end{aligned} \tag{2}$$

As defined by Ben-Tal and Nemirovski (1998), feasible solutions to (2) are *robust feasible solutions* and the optimal solution to (2) is a *robust optimal solution*. Bertsimas and Sim (2004) introduced the concept "price of robustness" which they define as "the effect to the objective function of the nominal problem" resulting from a decrease in the probability of constraint violation. In other words, how "heavily" is the objective function value penalized when we are guarded against violating constraints (Bertsimas and Sim, 2004)? Implicitly, this is the difference between the *robust optimal solution* and the objective function value of the nominal problem. In section 6, we explicitly define a similar measure called the cost of robustness.

Soyster (1973) was the first to show that uncertain linear programs could be formulated as robust convex linear programs such that feasibility was preserved for all possible values of the uncertain parameters defined within a set. Many authors, including Soyster, agreed that his formulation was overly conservative – too much of the optimal objective value was lost. In other words, robustness costs too much. Soyster's approach lay dormant until the early 1990s when Ben-Tal and Nemirovski reformulated a less conservative model, one that was more realistic for application. For example, in Soyster's model, the probability of violating the *i*th constraint was zero; in Ben-Tal's and Nemirovski's model (1998), the probability of violating the *i*th constraint was less than ε, with ε > 0. They showed the probability of constraint violation to be bounded above by $e^{-\Omega^2/2}$, where Ω is a user defined parameter which measures the trade-off between robustness and optimality. If Ω = 0, there is no robustness (the formulation is simply the nominal problem). As Ω increases, the probability of constraint violation decreases and the optimal objective value deteriorates. A drawback to their approach was that the robust counterpart was more difficult to solve than the nominal problem. For example, LPs become second order cone programs (SOCPs), SOCPs become semi-definite programs (SDPs) and SDPs are NP-hard.

Bertsimas and Sim (2004) suggested an alternative robust counterpart that retained the degree of the original problem – LPs stayed LPs. It was also less conservative than Soyster's model and, like Ben-Tal and Nemirovski's model, allowed the user to control the following: 1) the level of protection against the *i*th constraint being violated, 2) the probability the solution would be optimal and 3) the optimal value of the solution with respect to the nominal problem. In addition, they guaranteed that the solution is feasible given that no more than Γ uncertain parameters changed, where Γ is a user defined parameter. However, Bertsimas and Sim show that there is a probabilistic guarantee that decisions will remain feasible and that the robust optimal objective will be achieved or exceeded, even if more than Γ parameters change.

Bertsimas and Sim (2004) differed from Ben-Tal and Nemirovski (1998) in how uncertainty sets are defined, both in structure and scale, as well as in how their models allowed for the relationship between the level of robustness and optimality to be adjusted. Ben-Tal and Nemirovski modelled uncertainty sets by ellipsoids and intersections of ellipsoids, which increased the complexity of the problem, while Bertsimas and Sim modelled uncertainty by polyhedral sets which preserved the structure of the problem; the robust counterpart of an LP remained an LP. For both models, $a_i \in A$, $\forall i$, were unknown, mutually independent values symmetrically distributed with respect to $\bar{a}_i$ on the interval $[\bar{a}_i - \hat{a}_i, \bar{a}_i + \hat{a}_i]$.

Ellipsoidal uncertainty sets (Ben-Tal and Nemirovski, 1998) are given by

$$U^{\Omega} = \{a \in R^n : \sum \frac{(a_i - \bar{a}_i)^2}{\hat{a}_i^{\,2}} \le \Omega^2\}, \tag{3}$$

where Ω is the user defined parameter previously mentioned, which adjusts the trade-off between robustness and optimality. As Ω increases, the upper bound on the probability of constraint violation, $e^{-\Omega^2/2}$, decreases (Table 1). Therefore, the scale of the uncertainty sets (or ellipsoids) and the probability of constraint violation is determined by the parameter Ω, which is dependent upon a user's risk preference.

Ben-Tal and Nemirovski (1998) suggested ellipsoidal uncertainty because ellipsoids can approximate more complicated uncertainty sets well, they can represent stochastic uncertainty sets deterministically, and they have a convenient mathematical structure. However, ellipsoidal uncertainty means that the robust counterpart of an LP becomes a SOCP, but the authors argue that this is not a problem because large SOCPs can be solved efficiently in polynomial time (Ben-Tal and Nemirovski, 1998).

Bertsimas and Sim (2004) modelled uncertainty by box uncertainty sets,

$$U = \{a : |a_i - \bar{a}_i| \le \hat{a}_i, \forall i\}. \tag{4}$$

Recall that Γ is a user defined parameter that adjusts the robustness of the model. Assuming that at most Γ parameters take their worst case value in the future, the probability of constraint violation is bounded above by $e^{-\Gamma^2/2|J|}$ (Table 1). Thus, as Γ increases, more protection is given; hence the solution is more robust. In contrast to the parameter Ω, introduced by Ben-Tal and Nemirovski, Γ does not affect the size of the uncertainty set, as seen by the inclusion of Ω in Equation 3 and the absence of Γ in Equation 4.

Soyster ('73)	Ben-Tal & Nemirovski ('98)	Bertsimas & Sim ('04)
Pr$(a'x^* > b) = 0$ For all possible *a*.	**Pr**$(a'x^* > b) < \varepsilon$ For all possible *a*, where $0 < \varepsilon < e^{-\Omega^2/2}$	**Pr**$(a'x^* > b) < \varepsilon$ For Γ a_i values, where $0 < \varepsilon < e^{-\Gamma^2/2\|J\|}$ Where $\|J\|$ is the cardinality of the subset of uncertain parameters a_i.

Table 1. Comparison of robust optimization models; x^* is the optimal solution vector.

3 Robust Optimization and Portfolio Selection

In the portfolio selection problem an investor chooses the proportion of capital to be invested in each of *N* assets such that a desired set of goals is achieved. For example, an investor may want to maximize return, minimise risk, or hold a certain number of assets. Typically, short selling is not permitted, hence the proportions of capital invested in each asset must sum to 1. The set of goals depends on the investor. However, the ultimate goal is to maximize total wealth while minimizing the risk of losses that may result from the investment. Risk comes from the uncertainty of asset returns, which are unknown, and stochastic in nature.

Markowitz (1952) presented the well-known expected value - variance (*E-V*) model for portfolio optimization. *E-V* optimization was based on the following assumptions: 1) investors seek to maximize expected return, but minimise variance, 2) solely maximizing expected return does not yield diversified portfolios and 3) diversified portfolios are favoured over non-diversified portfolios. Consider the following model which minimises variance for a specified level of return:

$$\text{Minimise} \quad \sum_{i=1}^{N}\sum_{j=1}^{N} w_i w_j \sigma_{ij}$$

$$\text{Subject to} \quad \sum_{i=1}^{N} w_i \mu_i \geq \textit{Target Return},$$

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

$$w_i \geq 0, \qquad \forall i = 1..N.$$

Where μ_i is the expected return on asset i, σ_{ij} is the covariance of assets i and j, and w_i is the fraction of total wealth invested in asset i. Alternatively, we may maximize expected return for a specified level of risk.

For both models, let E be expected portfolio return and V portfolio variance. If S represents the set of all possible (E,V) combinations it is reasonable to assume that an investor would only consider the subset of portfolios which Markowitz termed "efficient." An efficient portfolio is one that yields the highest expected return for a specific variance or has the least variance for a specific expected return.

The underlying assumption of Markowitz's model is that the required observation and analysis leading to the probability beliefs concerning the uncertain parameters μ_i and σ_{ij} have been accomplished. *E-V* optimization applies those beliefs to achieve a desired portfolio. As seen by the *E-V* model discussed above, μ_i and σ_{ij} are treated as known constants. However, the distributions of asset returns are uncertain. Therefore, as Bienstock (2007) suggested, μ_i and σ_{ij} are not known constants, but soft quantities. It is reasonable to conclude that a model which treats returns as known constants will produce portfolios that underperform the optimal objective when returns are less than their expected values.

E-V optimization has encountered other criticisms with respect to the composition of efficient portfolios. Michaud (1989) suggested that *E-V* portfolios "don't make investment sense" because they maximize estimation error; efficient portfolios overweight assets with large expected returns and small variances, which are negatively correlated. These, Michaud argued, are the assets whose μ_i and σ_{ij} estimates are more likely to be furthest from their true values. Ceria and Stubbs (2006) agreed, suggesting that *E-V* optimization is "counterintuitive" and too sensitive to fluctuations in the first and second moments of asset returns. Hence, small changes in μ_i and σ_{ij} yield very different efficient (E,V) combinations. Ceria and Stubbs suggested that there are two approaches to overcome estimation error in *E-V* optimization: better estimation techniques or better techniques for optimization under uncertainty (2006). The authors argued that even though there are estimation techniques for μ_i and σ_{ij} that may produce a more stable *E-V* portfolio, statistical methods are driven by underlying distributions which is problematic. Therefore, uncertainty should be incorporated into the optimization process, creating a need for methods such as robust optimization that treat uncertain parameters as soft quantities in the optimization process, i.e. instead of using a single value such as μ_i, asset returns can take on any value within a defined set of possible outcomes.

As previously mentioned, the main reason for applying a robust optimisation framework to the portfolio selection problem is because asset returns are unknown, variable, and their distributions are uncertain. Although the distributions of asset returns are uncertain, we may assert that μ or σ, or both, belong to an uncertainty set, the bounds of which we can define. Most robust portfolio models describe asset returns by ellipsoidal uncertainty sets, based on the methodology of Ben-Tal and Nemirovski (1998) and El Ghaoui et al. (1997), in which the user defined parameter Ω adjusts the guaranteed and achieved robustness of the portfolio. Previously, robustness has been evaluated based upon performance, particularly worst-case performance, then compared to the worst-case performance of a non-robust model such as the *E-V* model. In addition to worst-case performance, we suggest that it is also important to evaluate robustness based upon whether a model yields portfolios that achieve their guaranteed robustness in practice (see section 6).

In 2000, Lobo & Boyd presented two robust portfolio models: the first, gave an upper-bound on the risk associated with a portfolio, given a set of decisions; the second, minimised the upper-bound on risk. They presented several different methods for modelling the uncertainty sets for the expected returns vector and covariance matrices, such as box or ellipsoidal sets. Each robust model was a semi-definite program solved via interior point methods. Their results focused on the performance of the solution method rather than on the robustness of the optimal portfolios.

Goldfarb and Iyengar (2001) defined asset returns by robust factor models in which uncertainty was modelled by ellipsoidal sets. The robustness of a robust Sharpe Ratio problem was evaluated based on performance, particularly in worst-case scenarios, and compared to the *E-V* portfolio model. Results showed the worst-case performance of the robust model was approximately 200% better than the non-robust model; thus, they concluded that robust portfolios were more apt to withstand noisy data.

El Ghaoui et al. (2003), introduced and evaluated the robustness of a worst-case VaR model in which the uncertainty (of both μ and σ) was modelled by ellipsoidal sets. Results showed that the non-robust model 'wins' if there is no uncertainty, but the robust model 'wins' in the worst-case scenario, which is what one would expect.

Tutuncu and Koenig (2004) describe μ and σ by uncertainty sets in order to optimize a model which seeks to find the "best worst-case" portfolio and compare its performance to *E-V* portfolios. Results showed that robust portfolios are only "marginally inefficient" when returns take their expected value but *E-V* portfolios are "severely inefficient" when returns take their worst-case values (as defined by their uncertainty set). In addition, robust portfolios tended to concentrate on a small set of asset classes, having chosen mostly large capital value stocks.

Ceria and Stubbs (2006) presented a model which minimised the difference between the estimated and actual efficient frontiers while maximising portfolio return. Typically the true frontier lies between the estimated and actual frontiers, hence, minimising their distance will bring them closer to the true frontier. Results showed that the robust model yielded greater "realized returns" in most cases.

Kim and Boyd (2007) presented the robust efficient frontier analysis method to address the problem of poor performance by *E-V* optimisation resulting from the use of estimates of μ and σ. Their main focus was to construct a worst-case efficient frontier representing the optimal trade-off between worst-case risk-return pairs in which the uncertainty in μ and σ are independent. They analyse the basic properties of a worst-case robust efficient frontier and present several "computationally tractable uncertainty models".

Pflug and Wozabal (2007) took a slightly different perspective on modelling uncertainty by considering the probability model of an asset to be unknown. They constructed a robust portfolio selection problem in which a 'confidence set' described the probability distribution of asset returns. In addition, they evaluated the trade-off between risk, robustness, and portfolio return. Their results showed that as robustness increases, risk and portfolio return decreases and portfolios were more diversified.

Robust optimization techniques have been criticized for giving equal weight to all possible values of the uncertain parameters, specified within their respective uncertainty sets, which may not be a realistic assumption (Bienstock, 2007). Bienstock (2007) addresses this criticism by defining two types of uncertainty sets that give higher weight to more significant data realizations. The first type defines returns based on a histogram of shortfalls from a point estimate, and the second type models correlations among deviations. Cutting-plane algorithms were introduced to solve the robust models resulting from both types of uncertainty sets.

More recently, Bertsimas and Pachamonova (2008) suggested a multi-period portfolio optimisation model built upon the approach of Ben-Tal et al (1999), but with polyhedral, instead of ellipsoidal, uncertainty sets. They compared the computational performance of their linear robust models with a single period mean-variance model using simulations of future returns of 3 assets. Results suggested that a robust multi-period approach should be considered as an alternative to single period *E-V* models.

As is evident from the literature, nearly all robust portfolio models construct uncertainty sets as ellipsoids, based on the work of Ben-tal and Nemirovski (1998) and El Ghaoui et al.(1997). Typically, solution robustness is evaluated by comparing the worst-case performance of the robust model with that of a non-robust model. In addition, there is not an explicit evaluation of the cost of robustness. In this paper, we consider the robust portfolio model of Bertsimas and Sim (2004), which constructs polyhedral uncertainty sets; we investigate the guaranteed and achieved robustness of the solution as well as the cost of robustness. Furthermore, by altering model parameters, we evaluate the robustness of the model itself, that is, as model parameters change, how much do our decisions change?

4 Linear Robust Counterpart to Portfolio Optimization

In this section, we discuss the formulation of the linear robust counterpart to the portfolio optimization problem, first from a duality perspective (introduced by Bertsimas and Sim (2004)) and we then explain the rationale for the model.

4.1 Robust Counterpart by Duality

The basic portfolio optimization problem is defined as follows:

$$\text{Max } \sum_i r_i w_i \,. \tag{5}$$
$$\text{S.t. } \sum_i w_i \le 1\,,$$
$$w_i \ge 0\,, \qquad \forall i\,.$$

Asset returns, r_i, are uncertain parameters with unknown distributions defined as bounded and symmetric with respect to a point estimate $\bar{r}_i$:

$$r_i \in [\bar{r}_i - \hat{r}_i, \bar{r}_i + \hat{r}_i]\,. \tag{6}$$

Even though the true distribution of r_i is unknown, historical data can be used to estimate the mean log return of asset i, and substituted for the point estimate $\bar{r}_i$. A new stochastic variable η_i (Bertsimas and Thiele, 2006) measures the deviation of parameter r_i from $\bar{r}_i$ and takes values in [-1,1],

$$\eta_i = \frac{r_i - \bar{r}_i}{\hat{r}_i}\,.$$

By rearranging this equation, r_i can be expressed as:

$$r_i = \bar{r}_i + \hat{r}_i \eta_i\,. \tag{7}$$

Let $|J|$ be the number of parameters, r_i, that are uncertain; then for Soyster's and Ben-Tal & Nemirovski's model

$$\sum_i \frac{|r_i - \bar{r}_i|}{\hat{r}_i} = |J| \quad \text{or} \quad \sum_i |\eta_i| = |J|\,.$$

Bertsimas and Sim (2004) relaxed this condition by defining a new parameter Γ (the *budget of uncertainty*) as the number of uncertain parameters that take their worst-case value $\bar{r}_i - \hat{r}_i$. Therefore,

$$\sum_i |\eta_i| \le \Gamma\,, \text{ such that } \Gamma \in [0, |J|]\,.$$

Rewriting the initial portfolio optimization problem using (7) for r_i:

$$\underset{w_i}{Max} \; (\sum_i \bar{r}_i w_i + \underset{\eta_i}{Min} \sum_i \hat{r}_i \eta_i w_i\,). \qquad \equiv \qquad \underset{w_i}{Max} \; (\sum_i \bar{r}_i w_i - \underset{\eta_i}{Max} \sum_i \hat{r}_i \eta_i w_i\,)\,.$$
$$\text{S.t. } \sum_i w_i \le 1\,, \qquad \qquad \text{S.t. } \sum_i w_i \le 1\,,$$

$$\sum_i |\eta_i| \le \Gamma, \qquad\qquad \sum_i \eta_i \le \Gamma,$$
$$w_i \ge 0,\ -1 \le \eta_i \le 1, \quad \forall i. \qquad\qquad w_i \ge 0,\ 0 \le \eta_i \le 1, \quad \forall i.$$

By duality (Bertsimas and Sim, 2004), the inner maximization problem subject to the stochastic constraints becomes:

$$\begin{aligned}
\underset{p,q_i}{Min} \quad & \Gamma p + \textstyle\sum_i q_i \,. \\
\text{S.t.} \quad & p + q_i \ge \hat{r}_i w_i, \qquad \forall i, \\
& p \ge 0, \\
& q_i \ge 0, \qquad \forall i.
\end{aligned}$$

Substituting this result, we obtain the following robust counterpart:

$$\begin{aligned}
\underset{w_i}{Max} \ & (\textstyle\sum_i \bar{r}_i w_i - \underset{p,q_i}{Min}(\Gamma p + \sum_i q_i)) & \equiv \quad & Max \ \textstyle\sum_i \bar{r}_i w_i - \Gamma p - \sum_i q_i \,. \\
\text{S.t.} \ & \textstyle\sum_i w_i \le 1, & & \text{S.t.} \ \textstyle\sum_i w_i \le 1, \\
& p + q_i \ge \hat{r}_i w_i, \quad \forall i, & & p + q_i \ge \hat{r}_i w_i, \quad \forall i, \\
& p \ge 0, & & p \ge 0, \\
& w_i,\ q_i \ge 0, \quad \forall i. & & w_i,\ q_i \ge 0, \quad \forall i.
\end{aligned}$$

4.2 Interpretation of Robust Counterpart

The robust counterpart of an uncertain optimization problem is a max-min or min-max model; the objective is to optimise the worst-case performance. Soyster's and Ben-Tal & Nemirovski's models stipulate that every constraint be feasible for every uncertain parameter defined within a bounded symmetric set (an uncertainty set). That is, their models are optimized for *every* uncertain parameter taking its worst-case value.

Bertsimas and Sim (2004) introduced a model that assumes at most Γ uncertain parameters will take their worst-case values, not *every* parameter. Applying this concept, consider the basic portfolio model stated in section 3.1, but using the definition of r_i in (7).

We desire the portfolio with the best worst-case return given that Γ asset returns take their worst-case values, $\bar{r}_i - \hat{r}_i$. Therefore,

$$\begin{aligned}
& \text{Max (Min } \textstyle\sum_i \bar{r}_i w_i - \sum_{t\in T} \hat{r}_t w_t \text{).} & \equiv \quad & \text{Max} \textstyle\sum_i \bar{r}_i w_i - \text{Min(Max} \sum_{t\in T} \hat{r}_t w_t \text{).} \\
& \text{S.t. } \textstyle\sum_i w_i \le 1, & & \text{S.t.} \quad \textstyle\sum_i w_i \le 1,
\end{aligned}$$

where $I = \{i | 1 \le i \le N\}$, $T \subseteq I$, $|T| = \Gamma$ and $t \in T$ are the subset of Γ assets that take their worst-case values, $\bar{r}_t - \hat{r}_t$. The min-max term in the objective seeks to minimise the sum with respect to w_i and maximize the sum by choosing the Γ assets with the largest $\hat{r}_i w_i$ as the subset *T*. For the moment, assume that the quantity $\hat{r}_t w_t$ is the same for all *t*, and denoted as *p*. Then the term $\sum_{t \in T} \hat{r}_t w_t$ becomes Γ*p*. Now consider the possibility that $p_t \neq p$ for some *t*, where $p_t = \hat{r}_t w_t$ for all *t*. Because the term $\sum_{t \in T} \hat{r}_t w_t$ will be greater than or equal to zero at optimality, we will only consider the case when $p_t \geq p$. Therefore, the difference $p_t - p$, for all *t*, needs to be added onto Γ*p* and the quantity Min(Max $\sum_{t \in T} \hat{r}_t w_t$) becomes:

$$\text{Min}[\Gamma p + \sum_t (p_t - p)], \qquad \forall t = \{t | (p_t - p) \geq 0\} \tag{8}$$

We can restrict the difference $p_t - p$ to be greater than or equal to zero if we introduce a new variable q_t given by the equation:

$$q_t = \max(0, p_t - p). \tag{9}$$

The question now is: Which p_t is chosen as our *p* value? Recall that the min-max term in the objective seeks to maximize the quantity $\sum_{t \in T} \hat{r}_t w_t$ by choosing the Γ largest $\hat{r}_i w_i$ as the subset *T* and that the quantity $\sum_t q_t$ is greater than zero only when $p_t > p$. Thus, *p* is chosen as the smallest $\hat{r}_t w_t$, over all *t*, which means it is the Γ^{th} largest $\hat{r}_i w_i$, over all *i*.

The final portfolio optimization model becomes:

$$\text{Max } \sum_i \bar{r}_i w_i - \Gamma p - \sum_i q_i . \tag{10}$$

$$\text{S.t.} \quad \sum_i w_i \leq 1 .$$

Remark: $\sum_{i \in I} q_i$ can be substituted for $\sum_{t \in T} q_t$ because every $\hat{r}_i w_i$, where $i \notin T$, will be less than *p*. Therefore, $p_i - p$ will be less than zero and the corresponding q_i will be zero.

5 Robust Portfolio Model for Uncorrelated Asset Returns

5.1 The Model

Bertsimas and Sim (2004) reformulated a maximum expected return portfolio model as a linear robust optimization problem, as shown in section 4:

$$\text{Maximize} \qquad \sum_{i=1}^{n} \bar{r}_i w_i - p\Gamma - \sum_{i=1}^{n} q_i \tag{11}$$

Subject to $\sum_{i=1}^{n} w_i \leq 1$,

$$p + q_i \geq c\hat{r}_i w_i, \qquad \forall i,$$
$$w_i \geq 0, \qquad \forall i,$$
$$q_i, p \geq 0, \qquad \forall i,$$

where $\bar{r}_i$ is the point estimate for the log return of asset i (e.g. the median or mean log return), w_i is the proportion of total wealth invested in asset i and r_i is the true log return of asset i. The true log return of asset i, r_i, belongs to the interval $[\bar{r}_i - c\hat{r}_i, \bar{r}_i + c\hat{r}_i]$, where $\hat{r}_i$ is chosen by the user and determines how the uncertainty set defines r_i, and $c \in \Re^+$ defines the magnitude of the range of the set. For example, if $\hat{r}_i$ is the standard deviation of asset i, then c would determine the width of the interval in terms of the number of standard deviations. Alternatively, if $\hat{r}_i = \bar{r}_i$, where $\bar{r}_i$ is the mean log return, then c would determine the width of the interval in terms of the percentage of $\bar{r}_i$ that the true log return deviates from $\bar{r}_i$. The user defined parameter $\Gamma \in \Re^+$ is given a value between 0 and n, inclusive. Bertsimas and Sim (2004) allow Γ to take non-integer values, however, we restrict Γ to be integer for simplicity because we did not observe any significant benefit from allowing Γ to take fractional values; thus, $\Gamma \in Z^+$. As Γ increases, the probability of underperforming the robust optimal objective decreases. At optimality, p is the Γ^{th} largest $c\hat{r}_i w_i$ and $q_i = \max(0, c\hat{r}_i w_i - p)$ for each asset i. The focus of the Bertsimas and Sim paper was to present their robust approach and not portfolio optimization per se; their experimental results were for a set of 150 stocks with $\bar{r}_i$ and $\hat{r}_i$ generated by arithmetic progressions.

5.1.1 Computational Platform

We chose to optimise this particular linear robust portfolio model using the solver CPLEX version 10.1, a common computational platform, within the modelling language AMPL. More specialized modelling languages, such as CVX which is implemented in Matlab (Grant, 2008), may be chosen for solving such convex problems.

5.2 Portfolio Composition

Data Set. The data set used to construct Figure 1, Figure 2, and Figure 3, consist of the monthly returns of 30 stocks selected at random from the FTSE 100 index beginning 1 January 1992 through to 1 December 2002. Other investigations were completed using larger asset pools: 169 stocks selected from the FTSE250, 248 stocks selected from the FTSE350, and 441 assets selected from the S&P500, beginning 1 October 1998 through to 1 September 2008. The properties of the robust model which include diversification, the selection of assets, and the weighting of assets were observed to be the same for N = 30 and for these larger asset pools. We chose to illustrate these properties using the results from N = 30 to simplify illustrations.

Diversification. Consider $N + 1$ consecutive portfolios corresponding to integer values of Γ from 0 to N. When $\Gamma = 0$, the portfolio consists of 1 asset; this is simply the maximum return problem with no robustness. As Γ increases, the number of assets increases until either all N assets are included or a maximum number of assets is reached. The composition of portfolios for successive values of Γ remains constant until all but 1 asset are dropped, which corresponds to $\Gamma = \Gamma_{drop}$. This behaviour is shown in Figure 1.

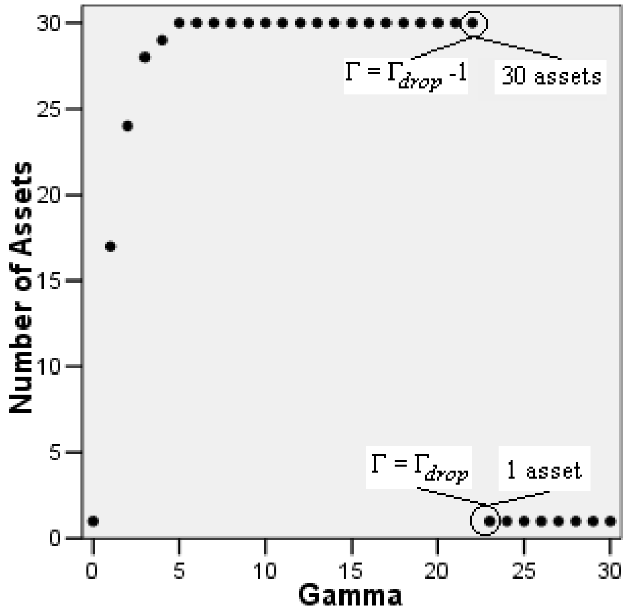


Figure 1. Number of assets selected for the portfolio at each Γ, when $N = 30$.

Selection & Weights. Again, consider $N + 1$ consecutive optimal portfolios corresponding to integer values of Γ from 0 to N. When $\Gamma = 0$, the portfolio consists of the asset with the largest $\bar{r}_i$. In Figure 2, the assets along the x-axis are in descending order by $\bar{r}_i$; asset 1 has the largest $\bar{r}_i$ and asset 30 has the smallest $\bar{r}_i$. Figure 2 has two purposes: 1) to show the assets held in each portfolio when $\Gamma = 0$ through to $\Gamma = 5$ and 2) to show how much the weight of each asset held at Γ decreases when Γ increases to $\Gamma + 1$. From Figure 2 we can see that as Γ increases and more assets are included, those with a larger $\bar{r}_i$ will be selected first and will also be in each subsequent portfolio until Γ_{drop} is reached. In Figure 3, the assets along the x-axis are in ascending order by $\hat{r}_i$; asset 1 has the smallest $\hat{r}_i$ and asset 30 has the largest $\hat{r}_i$. From Figure 3, we can see that once selected, the assets with the smallest $\hat{r}_i$ are given the most weight. In addition, Figure 2 shows that as Γ increases and more assets are selected, the weight of each asset that belongs to the previous portfolio decreases by the same percentage. Lastly, for $\Gamma \geq \Gamma_{drop}$ the optimal portfolio consists of the asset with the largest risk-adjusted return, $\bar{r}_i - c\hat{r}_i$ (not shown in Figure 2 or Figure 3).

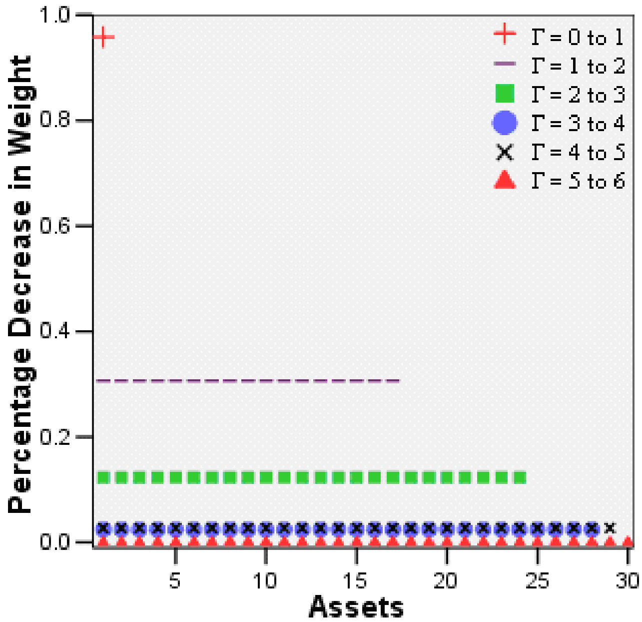


Figure 2. Assets in descending order by $\bar{r}_i$. An example of how assets are selected and how weights change as more assets are included in the portfolio.

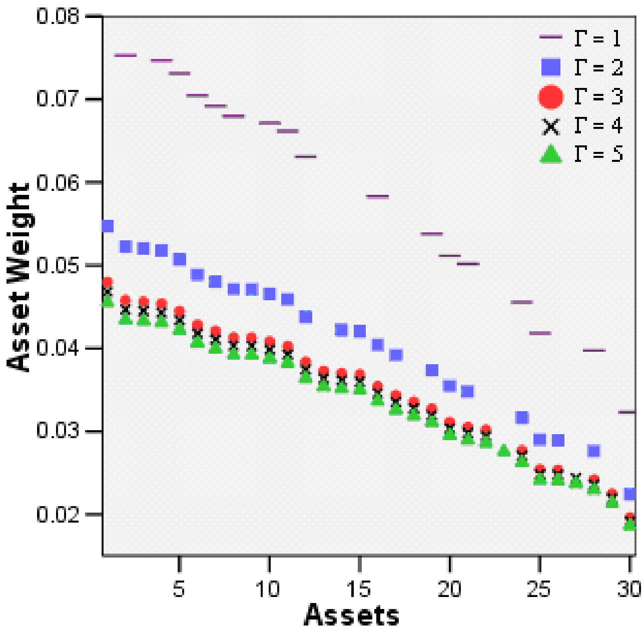


Figure 3. Assets in ascending order by $\hat{r}_i$. An example of how robust models weight assets.

6 The Cost of Robustness

Robustness, viewed as a performance guarantee, comes at a cost. In the case of portfolio optimization, it is the probability guarantee that the portfolio return will be at least equal to that of the robust optimal solution. One would expect that in order to achieve robustness, a sacrifice, in terms of optimal objective value, will occur. But how much does this sacrifice cost? And is it worth it? There is a two-fold motivation to the investigation detailed in this section. Firstly, to provide a measure for the cost of robustness and determine if the robust methodology in section 5 is robust. Secondly, how are the cost and robustness (both guaranteed and achieved robustness) affected when the following are changed: *i)* $\bar{r}_i$, $\hat{r}_i$, and c defining the uncertainty set of r_i and/or *ii)* the size of the set of historical data used to estimate $\bar{r}_i$ and $\hat{r}_i$.

The data set used in this investigation consists of 120 monthly log returns of 68 assets from the FTSE 100 starting 1 February 1996 through to 1 January 2007. A set of $m \in M$ months, where $M = \{20, 40, 60\}$, was randomly selected and used to generate 10 robust models, R_j, where j = 1..10 (Table 2), each with a different uncertainty set defining r_i, the true log return of asset i, which is unknown and variable. For each model, 100 instances (henceforth referred to as trials, t) were randomly generated in order to obtain a distribution for both measures of cost and both measures of robustness detailed in sections 6.1 and 6.2. Thus for each t, from 1 to 100, a set of m randomly selected months (which was different for each trial) was considered as the set of available historical data and used to optimise 10 robust models, for a total of 1000 optimal portfolios. The set of m months was also used for in-sample back-testing in the evaluation of robustness.

Model R_j	$\bar{r}_i$	$\hat{r}_i$	c
j = 1..3	Mean Log Return	Standard Deviation	1, 2, 3
j = 4..7	Mean Log Return	Mean Log Return	0.90, 0.95, 0.98, 1
j = 8..10	Median Log Return	Standard Deviation	1, 2,3

Table 2. Summary of robust models.

For each trial t, the m randomly selected months were used to estimate the value $\bar{r}_i$ and $\hat{r}_i$; each robust optimal solution, $Z^{Opt}_{j,t,m}$, was obtained using formulation (11) from section 5.1. The Γ value that yielded the most robust diversified portfolio (i.e. the portfolio with the smallest probability of underperformance and consisting of more than one asset) was chosen as the optimal robust portfolio for model R_j, for each trial t and each set of $m \in M$ months. A characteristic of the robust models, which was discussed previously in section 5.2, is that as Γ increases from an initial value of 0, the number of assets in each corresponding portfolio also increases until all but one asset are suddenly dropped, which corresponds to $\Gamma = \Gamma_{drop}$. Results show that when $\Gamma = \Gamma_{drop} - 1$, the resulting portfolio is the most robust diversified portfolio consisting of at least as many assets as each portfolio corresponding to all other values of Γ. This value of Γ and hence, the probability of underperformance (Bertsimas and Thiele, 2006) given by (12), may be different for many of the 1000 portfolios, for each set m. However, since there are only N = 68 possible values of Γ_{drop}, many of the portfolios will have the same probability of underperformance.

$$\Pr(Z^{true}_{j,t,m,l} \le Z^{Opt}_{j,t,m}) \le 1 - \Phi((\Gamma - 1)/\sqrt{N}), \quad (12)$$

where $Z^{true}_{j,t,m,l}$ is the realized portfolio return for model R_j, during trial t, given the set of m months, evaluated at month l such that l = 1..m. $Z^{Opt}_{j,t,m}$ is the robust optimal objective function value for Γ, and N is the total number of assets.

6.1 Measures of Cost

For each trial t and set m, robust model R_j yields a n-vector of optimal asset weights, $\mathbf{w}^{*}_{j,t,m}$. Therefore, the total portfolio return of R_j for each t and m, is given by $P^{Total}_{j,t,m}$ in Equation (13):

$$P^{Total}_{j,t,m} = \sum_{i=1}^{n} \bar{r}_{i,t,m} w^{*}_{i,j,t,m}, \quad \forall j,t,m, \tag{13}$$

where $\bar{r}_{i,t,m}$ is the mean log return of asset i over a set of m months for trial t. We introduce two measures for the cost of robustness. Let $r^{MMax}_{t,m}$ denote the return of the asset with the largest mean log return for trial t and set of m months. $Cost1$ and $Cost2$ measure the cost of the robust optimal portfolio, $P^{Total}_{j,t,m}$, with respect to $r^{MMax}_{t,m}$. $Cost1$ measures the deviation between the value of the non-robust solution (i.e. with just a single asset) and the value of the robust solutions, whereas $Cost2$ measures the percentage deviation.

$$Cost1_{j,t,m} = r^{MMax}_{t,m} - P^{Total}_{j,t,m}, \qquad \forall j,t,m. \tag{14}$$

$$Cost2_{j,t,m} = (r^{MMax}_{t,m} - P^{Total}_{j,t,m}) / r^{MMax}_{t,m}, \qquad \forall j,t,m. \tag{15}$$

6.1.1 Measures of Cost: Results

As discussed in section 6, the motivation for the cost analysis is to determine 1) the distributions of the cost of $P^{Total}_{j,t,m}$ for each model and 2) observe how changes in model parameters $\bar{r}_i$, $\hat{r}_i$, and c and/or the size of the data set $m \in M$, where M = {20, 40, 60}, affects cost.
Figure 4 and
Figure **5** show the distributions of $Cost1$ for each R_j, when m is 20 and 60 months, respectively. For each model, the minimum, maximum, median, and mean costs are plotted. In each Figure, each of the sets of piecewise linear functions corresponds to different values of c for fixed values of $\bar{r}_i$ and $\hat{r}_i$. Thus we can observe the effects of a change in c on each statistic and distribution. For example, consider models R_1, R_2, and R_3 for m = 20 months (
Figure 4). As c increases from 1 to 2 (R_1 to R_2), the value of each statistic of $Cost1$ also increases; as c increases from 2 to 3 (R_2 to R_3), the value of each statistic of $Cost1$ does not change. Similar behaviour is observed between models R_8, R_9, and R_{10}. Models R_4 to R_7, show an increase in the value of each statistic of $Cost1$ corresponding to an increase in c, particularly noticeable in maximum cost (which is an outlier). Moreover, the histograms of $Cost1$, for m = 20, show that the distribution of each R_j is close to bell-shaped, but positively skewed with slightly higher peaks and

1 outlier (the maximum). An example of these characteristics is given in Figure 6, which shows the histograms of *Cost*1 for models R_1, R_6, and R_{10}. Lastly,
Figure **4** shows that, as *c* increases, the distribution of *Cost*1 maintains a similar shape but is shifted upward. In other words, as the magnitude of the range of the uncertainty set increases, the cost of robustness also tends to increase.

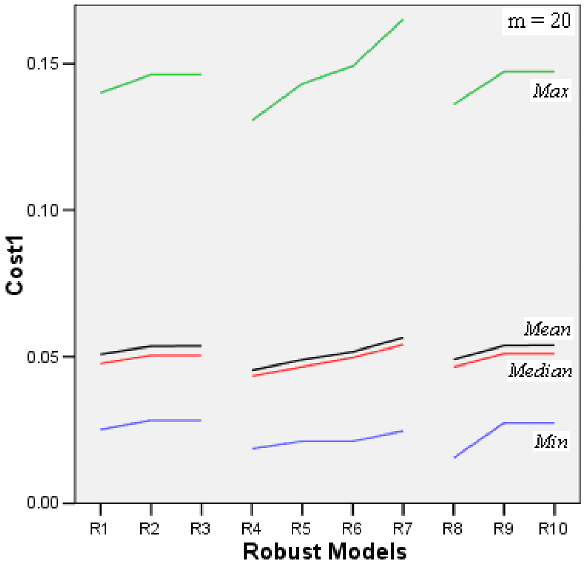


Figure 4. Distributions of *Cost*1 for each model R_j, when m = 20 months.

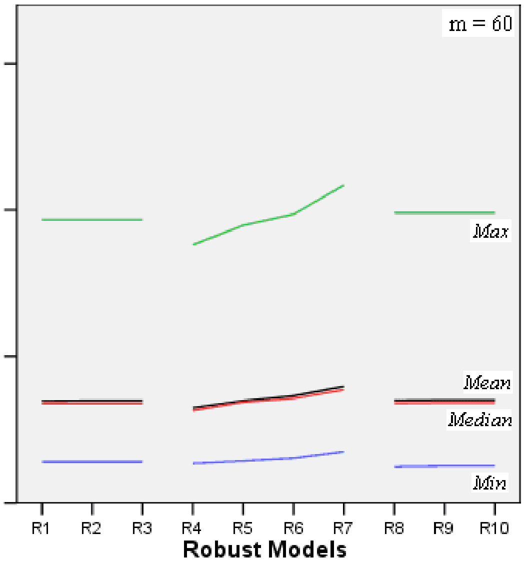


Figure 5. Distributions of *Cost*1 for each model R_j, when m = 60 months.

Figure **5** shows that when m = 60, an increase in *c* does not affect the min, max, median, or mean of *Cost*1 for R_1 to R_3 or R_8 to R_{10}; however, as when m = 20, models R_4 to R_7 show an increase in the value of each statistic of *Cost1* corresponding to an increase in *c*, and again, this is particularly noticeable for the maximum cost (an outlier). In addition, the distributions of *Cost*1 are close to bell-shaped, but positively skewed with slightly higher peaks and either 2 or 3 outliers (the 2 or 3 largest values); this can be observed in Figure 7, which shows the histograms of *Cost*1 for models R_1, R_6, and R_{10}. Finally,
Figure **5** shows that when m = 60, in contrast to m = 20, an increase in *c* has a minuscule effect on cost for models R_1 to R_3 and R_8 to R_{10}. Therefore, increasing the range of the uncertainty set does not significantly increase the cost of robustness for these 6 models, however, the cost of robustness for models R_4 to R_7 does tend to increase.

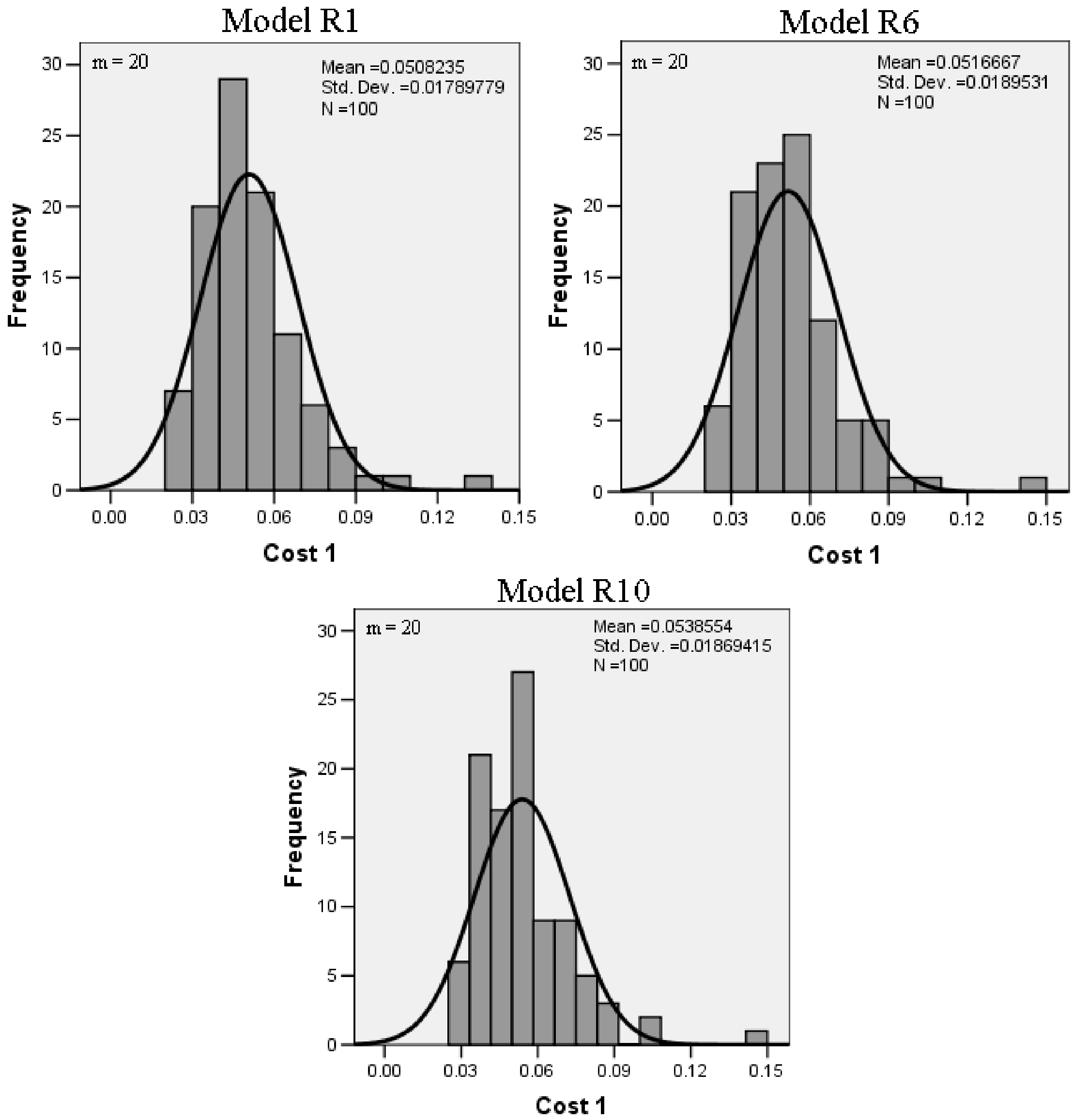
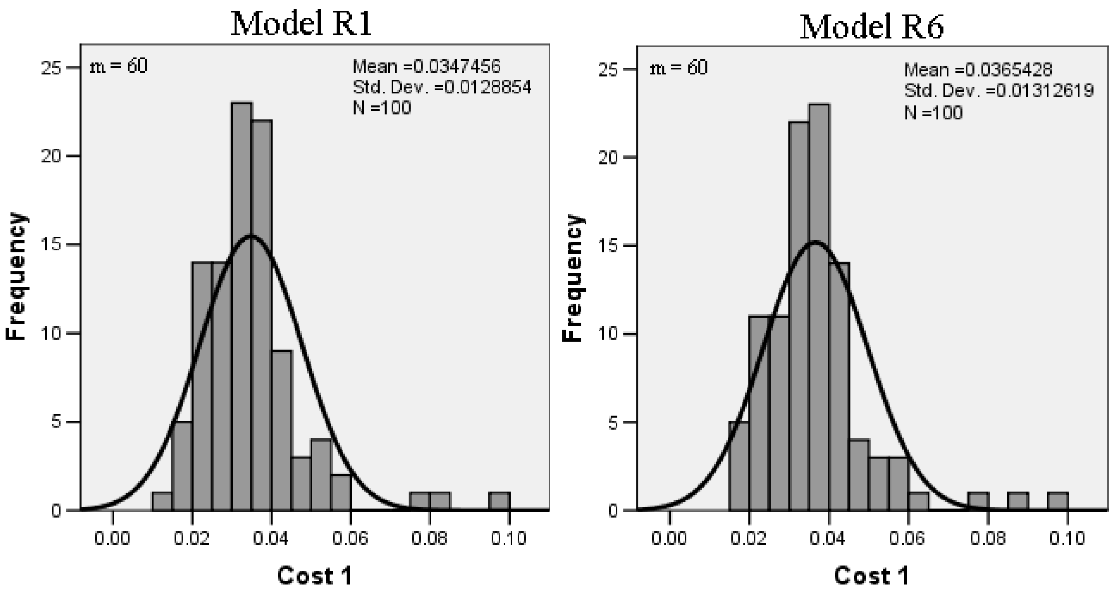


Figure 6. Histograms of *Cost*1 for Models R_1, R_6, and R_{10}, where m = 20 months.



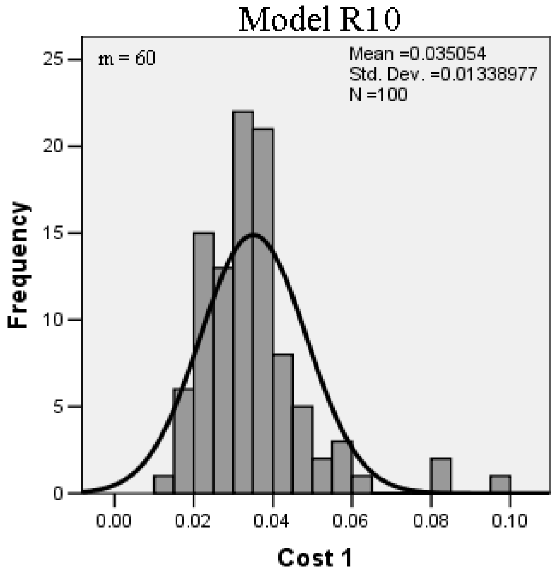


Figure 7. Histograms of *Cost*1 for Models R_1, R_6, and R_{10}, where m = 60 months.

As seen in Figure **4** and Figure **5**, and mentioned before, there are instances in which an increase in c does not appear to affect the distribution of *Cost*1. For m = 20 months, the composition of portfolios for c = 2 and c = 3 reveal that the distributions of the corresponding portfolios, R_2 & R_3 and R_9 & R_{10}, are so similar because in many of the trials the optimal weights were either identical or very similar. For example, both R_2 and R_3 yielded the same decisions for 91 out of 100 trials. The same is true when m = 60 months, but for c = 1..3.

From Figure **4** and Figure **5** we can also observe how a change in $\bar{r}_i$, $\hat{r}_i$, and m affects each statistic and distribution. For example, compare the distributions of R_1, R_2, and R_3 with those of R_4, R_5, R_6, and R_7. The latter set of models (with the exception of R_7), in which both $\bar{r}_i$ and $\hat{r}_i$ are defined as the mean log return of asset i, tend to cost less than the former as their distributions are shifted down; however, they tend to have larger outliers (maxima). A similar comparison can be made between the distributions of R_8 to R_{10} and R_4 to R_7. Now compare the distributions of *Cost*1 when m = 20 and m = 60 months, for R_1, R_2, and R_3; an increase in the number of months of historical data resulted in lower costs. There are several possible causes: 1) increasing the size of the dataset results in more precise estimates of $\bar{r}_i$ and $\hat{r}_i$, which yield more cost effective solutions, 2) $r_{t,m}^{MMax}$ may be significantly larger for each trial when m = 20, or 3) a combination of both 1) and 2). Results for m = 40 months (not shown) suggest that it is the first cause, because the distributions for *Cost*1 fall between those for 20 and 60 months. *Cost*2, shown in Figure 8 and Figure 9, provides a more accurate indication of the effect of changing m.

The distributions of *Cost*2 for each R_j, when m is 20 and 60 months, are shown in

Figure 8 and Figure 9, respectively. As with *Cost*1, the minimum, maximum, median, and mean are plotted for each model. In each Figure, each set of piecewise linear functions corresponds to different values of *c* for fixed values of $\bar{r}_i$ and $\hat{r}_i$. From Figure 8 and Figure 9 we can observe how a change in $\bar{r}_i$, $\hat{r}_i$, m and c affects each statistic and distribution. As in the case of *Cost*1, an increase in *c* tends to correspond to an increase *Cost*2, and, with the exception of the maximum for R_1 to R_3, each model tends to cost less when $m = 60$ than when $m = 20$, although only slightly. One distinction from *Cost*1 is that models R_4 to R_7 tend to cost less than the other 6 models, as seen by their distributions and histograms; they have means close to the other 6 models, but a much smaller spread, and no outliers. The histograms ($m = 20$, 60) for each R_j, $j = 1..3, 8..10$, are close to bell-shaped, but positively skewed with higher peaks. Figure 10 shows the histograms of *Cost*2 for models R_3, R_4, and R_{10}, when $m = 20$ months and Figure 11 shows the histograms of *Cost*2 for models R_3, R_6, and R_{10}, when $m = 60$ months. Models R_1 to R_3 have 1 or 2 outliers which are maxima whereas R_8 has an outlier that is a minimum, but only for $m = 20$ months. From an analysis of Figure 8 and Figure 9 and the 99% confidence interval for the mean of each model, we conclude that one can expect models R_1 to R_3 and R_8 to R_{10} to cost approximately 72-90% on average, models R_4 to R_6 to cost approximately 68-82% on average, and model R_7 to cost approximately 83-90% on average when the number of months in the historical data set is between 20 and 60 months. It is possible that further increasing the number of months in the historical data set will result in decreased costs, however, further investigations not presented in this paper suggest that the mean of *Cost*2 will not decrease significantly.

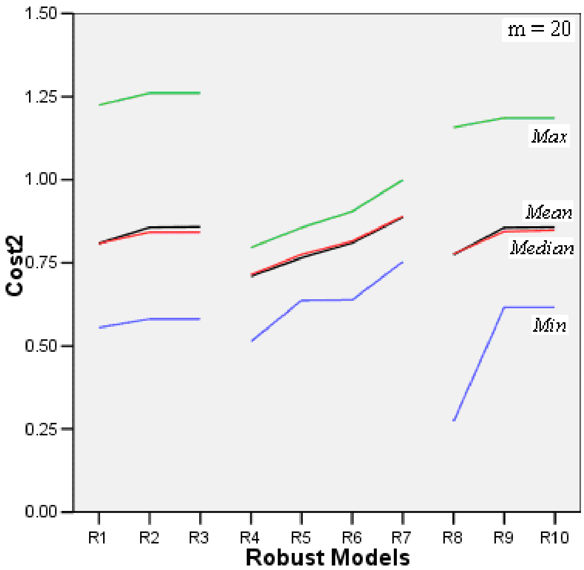


Figure 8. Distributions of *Cost*2 for each model R_j, when $m = 20$ months.

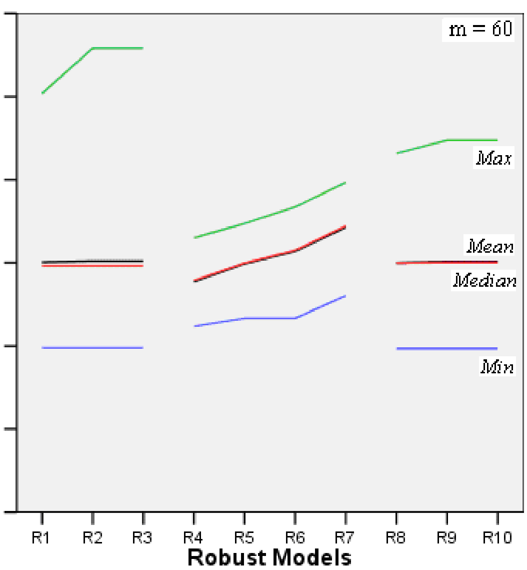


Figure 9. Distributions of *Cost*2 for each model R_j, when $m = 60$ months.

To summarise, the distribution of the loss in portfolio return with respect to $r_{t,m}^{MMax}$ for all ten models is very similar, with none having a significant advantage over the others, particularly with respect to the mean and median of both *Cost*1 and *Cost*2. However, the distributions of the percentage loss in portfolio return with respect to $r_{t,m}^{MMax}$ show models R_4 to R_7 to have a more consistent percentage loss, in terms of the spread of the distribution. The mean percentage loss for each model, regardless of $\bar{r}_i$, $\hat{r}_i$, m or c, is approximately 70-85%. Also, the results suggest that increasing the range of the uncertainty set defining r_i tends to increase costs whereas increasing the number of months in the set of historical data tends to decrease costs.

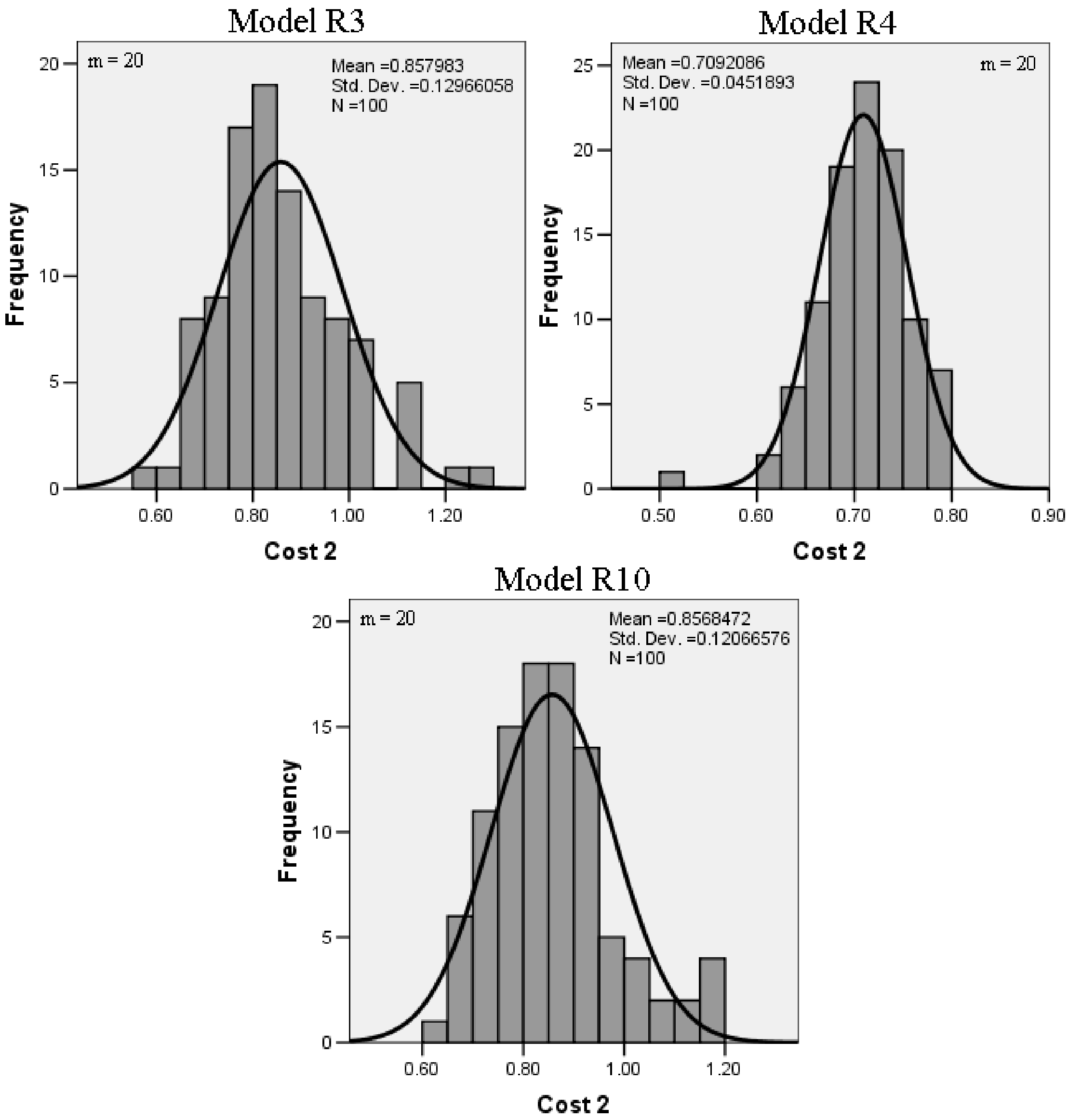


Figure 10. Histograms of *Cost*2 for Models R_3, R_4, and R_{10}, where m = 20 months.

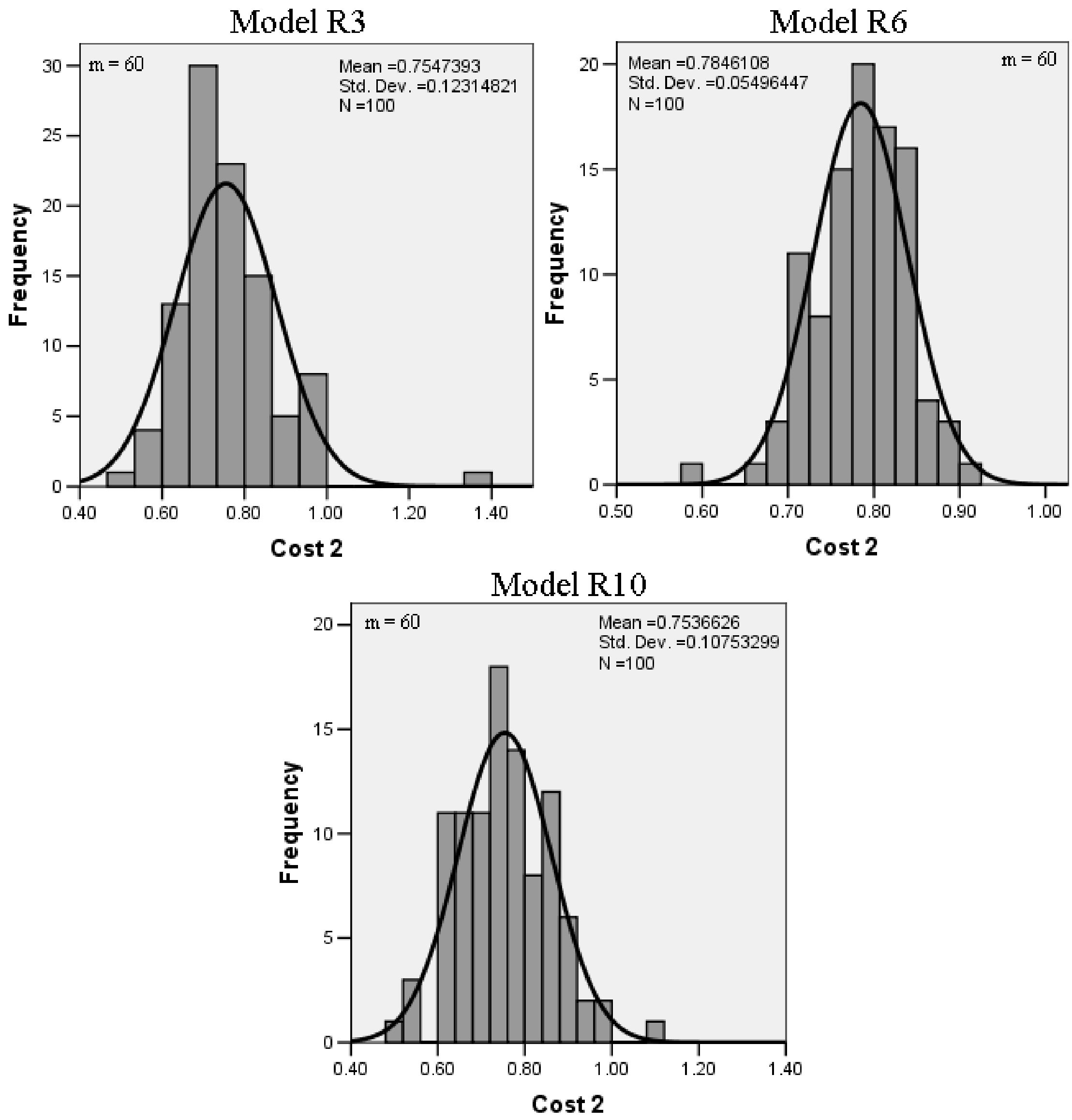


Figure 11. Histograms of *Cost*2 for Models R_3, R_6, and R_{10}, where m = 60 months.

6.2 Measures of Robustness

Robustness is assessed by 1) the probability of underperformance (which is dependent upon Γ) and 2) the proportion of evaluated portfolios that underperform the robust optimal objective (*PLO*: Proportion of portfolios Less than Objective), given by (18) and (19) respectively.

$$PLO^{Max}_{j,t,m} = (1-\Pr(Z^{true}_{j,t,m,l} \geq Z^{Opt}_{j,t,m})) = (1-\Phi((\Gamma-1)/\sqrt{N})), \qquad \forall j,t,m. \qquad (16)$$

$$PLO^{Eval}_{j,t,m} = \sum\nolimits_l \delta_{j,t,m,l} \,/\, m, \qquad \forall j,t,m, \qquad (17)$$

where $\delta_{j,t,m,l}$ is 1 if $Z^{true}_{j,t,m,l} < Z^{Opt}_{j,t,m}$ and 0 otherwise. Because both (16) and (17) are measures of underperformance, as they decrease, robustness increases and as they increase, robustness decreases. Comparing the distributions of $PLO_{j,t,m}^{Max}$ and $PLO_{j,t,m}^{Eval}$, $\forall j, m$, we can evaluate the robustness of this methodology for the stated definitions of r_i, c, and m.

6.2.1 Measures of Robustness: Results

As previously mentioned, the motivation of the robust analysis is to determine 1) if the robustness guaranteed by each model is actually achieved and 2) how the robustness of the solution (guaranteed and achieved) is affected by changes in model parameters $\bar{r}_i$, $\hat{r}_i$, and c or the size of the data set $m \in M$, where $M = \{20, 40, 60\}$.

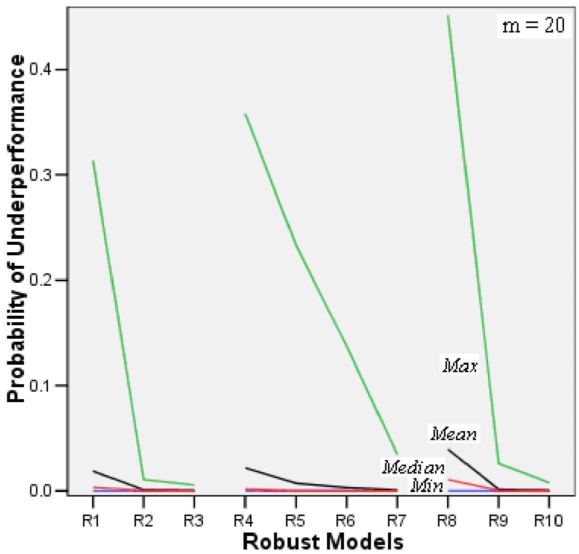


Figure 12. Distributions of $PLO_{j,t,m}^{Max}$ for each model R_j, when $m = 20$.

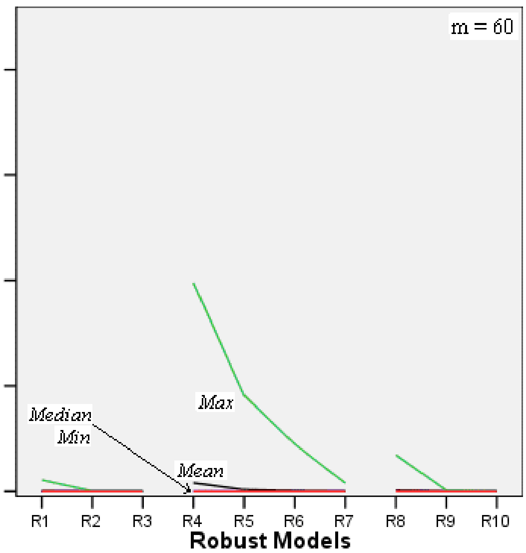


Figure 13. Distributions of $PLO_{j,t,m}^{Max}$ for each model R_j, when $m = 60$.

The guaranteed robustness of each model for $m = 20$ and $m = 60$ months is shown in Figure 12 and Figure 13, respectively. For each model, the minimum, maximum, median, and mean probability of underperformance are plotted. In each Figure, each set of piecewise linear functions corresponds to different values of c for fixed values of $\bar{r}_i$ and $\hat{r}_i$. Figure 12 and Figure 13 show that for both 20 and 60 months, an increase in c decreases the probability of underperformance; the distributions become tighter and means and medians are closer to 0. In addition, the plots suggest that an increase in m tends to result in a decrease in the probability of underperformance, thus, greater guaranteed robustness. For example, the maximum for each model in Figure 12 is much higher than the maximum in Figure 13, i.e. R_1 has a max close to 0.30 when $m = 20$ but a max of about 0.02 when $m = 60$. In addition, the distribution of each model, when m is 60 months, has a much smaller spread and a mean and median closer to 0. Recall that the probability of underperformance, $PLO_{j,t,m}^{Max}$, is dependent upon $\Gamma_{drop} - 1$; the larger $\Gamma_{drop} - 1$, the smaller $PLO_{j,t,m}^{Max}$ will be. Therefore, a larger data set of m months yields diversified portfolios for larger values of Γ, thus guaranteeing greater robustness. Lastly, although there is not one type of model that guarantees significantly more robustness, those which define $\bar{r}_i$ as the mean log return and $\hat{r}_i$ as the standard deviation of asset i, appear to be slightly more advantageous for both 20 and 60 months, as seen by tighter distributions and smaller maximum values in Figure 12 and Figure 13.

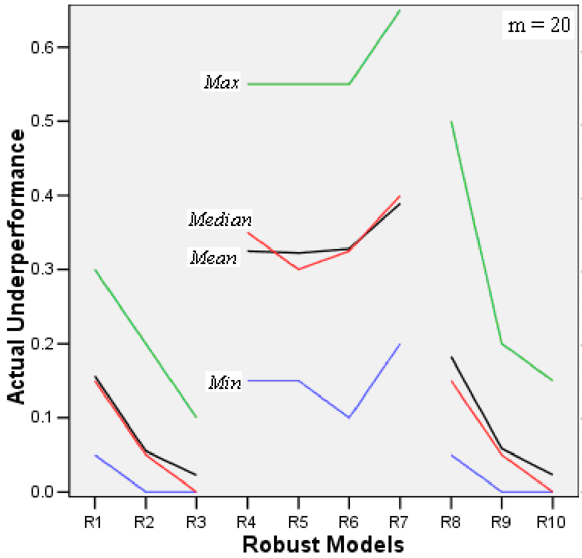


Figure 14. Distributions of $PLO_{j,t,m}^{Eval}$ for each model R_j, when m = 20 Mos.

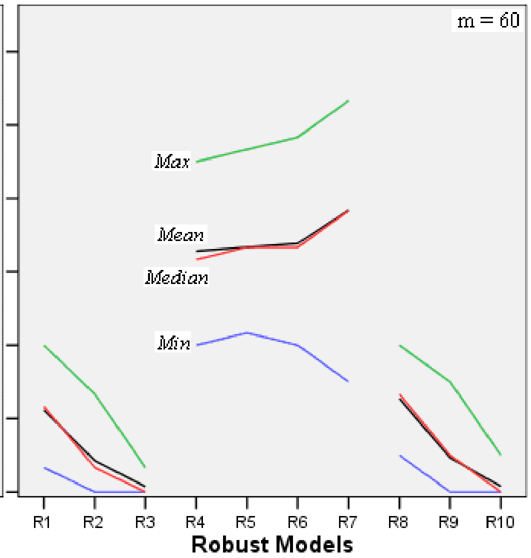


Figure 15. Distributions of $PLO_{j,t,m}^{Eval}$ for each model R_j, when m = 60 Mos.

The proportion of portfolios that actually underperform the robust optimal objective, for each model given m equals 20 or 60 months, is shown in Figure 14 and Figure 15, respectively. For each R_j, the minimum, maximum, median, and mean underperformance are plotted. In each Figure, each set of piecewise linear functions corresponds to different values of c for fixed values of $\bar{r}_i$ and $\hat{r}_i$. Both Figure 14 and Figure 15 show that increasing c also increases robustness when $\bar{r}_i$ is defined as the mean or median log return and $\hat{r}_i$ as the standard deviation of asset i; this is also shown in Table 3. However, when both $\bar{r}_i$ and $\hat{r}_i$ are defined as the mean log return (R_4 to R_7), increasing c tends to decrease robustness, seen by an increase in the mean and median. Recall that an increase in c results in higher costs, therefore, as models R_4 to R_7 increase in cost they are less robust. Table 3 shows the number of trials in which the proportion of portfolios that underperformed was less than the probability of underperformance. R_3 and R_{10}, in which c = 3, have the largest percentage of trials that achieve or exceed the guaranteed level of robustness, i.e. in over 60% of the trials the percentage of portfolios that underperformed the robust optimal objective was less than the probability of underperformance. In addition, models R_4 to R_7, there was not one trial in which the guaranteed level of robustness was achieved or exceeded for any set m.

How often is the guaranteed robustness achieved?

m	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10
20	1	29	63	**0**	**0**	**0**	**0**	2	27	63
40	1	11	73	**0**	**0**	**0**	**0**	1	9	67
60	**0**	9	64	**0**	**0**	**0**	**0**	**0**	8	62

Table 3. The number of trials (out of 100) in which the guaranteed level of robustness was achieved or exceeded, for each model R_j for a given set of m months.

Lastly, Figure 14 and Figure 15 show that increasing m tends to decrease the probability of underperformance, as seen by decreased max values and tighter distributions for the same model. However, as Table 3 shows, this does not necessarily correspond to an increase in the number of trials that achieve or exceed the guaranteed robustness. As we saw in Figure 12 and Figure 13 an increase in m also increases guaranteed robustness, which in most cases is less than 1%. In order for the actual percentage of portfolios that underperform to be less than 1%, for any given trial, the portfolio return for every month l (l = 1..m) must be greater than the robust optimal objective – not one out of the m months can underperform. For models R_2, R_3, R_9, and R_{10}, when m = 40 or 60, many trials did not achieve guaranteed robustness because only 1 portfolio out of the 40 or 60 underperformed the robust optimal objective.

In summary, the distributions of the probability of underperformance suggest that increasing the range of the uncertainty set defining r_i decreases the probability that the actual portfolio return will underperform the robust optimal objective. When $\bar{r}_i$ is the mean or median log return and $\hat{r}_i$ is the standard deviation of asset i, the actual proportion of portfolios that underperform also decreases, i.e. they are more robust. However when both $\bar{r}_i$ and $\hat{r}_i$ are the mean log return of asset i, the actual proportion of portfolios that underperform tends to increase, which means portfolios are less robust than they are guaranteed to be and they are much less robust than the other 6 models.

6.3 Discussion of Cost and Robustness Results

When $\bar{r}_i$ and $\hat{r}_i$ are both specified as the mean log return of asset i (R_4 to R_7), portfolios are slightly less costly, with respect to *Cost*1 and *Cost*2, but also less robust than the other models, particularly with respect to achieved robustness. In addition, an increase in the range of the uncertainty sets not only increases cost but decreases robustness, which is counterintuitive. One would expect that in exchange for a sacrifice in portfolio return there would be an increase in achieved robustness, i.e. fewer portfolios would underperform the optimal objective function value. For the other six models, increasing the range of the uncertainty set also increases cost, but that is in exchange for increased robustness. The results suggest that the most robust models are those which define $\bar{r}_i$ as the mean or median log return and $\hat{r}_i$ as the standard deviation of asset i.

7 Summary and Conclusions

In this paper we derived the robust optimization methodology of Bertsimas and Sim (2004) from a min-regret point of view. Robust optimization is best applied in situations where parameter values are unknown, variable, and their distributions are uncertain. In the case when distributions can be precisely estimated one should consider other methodologies such as stochastic programming. We have shown that the composition of robust portfolios are intuitive in nature because assets are first selected for the portfolio in descending order by mean log return and then those with smaller $\hat{r}_i$ are given more weight. In addition, robust portfolios are diversified both in terms of the number of assets and in weight, which is an advantageous feature, especially because portfolios are more diversified for larger values of Γ (through to $\Gamma_{drop} - 1$) which corresponds to greater guaranteed robustness.

The results of the investigation reported in this paper show that robust models in which $\bar{r}_i$ is the mean or median log return and $\hat{r}_i$ is the standard deviation of asset i, yield the most robust and cost effective portfolios. Furthermore, for these models, uncertainty sets with a larger range result in higher costs, but increased robustness. Results suggest that a value of $c \geq 2$ is preferred, i.e. r_i is defined as being within at least 2 standard deviations of its mean or median log return. The risk preferences of the investor determines the value of c chosen. In exchange for increased guaranteed and achieved robustness, the 99% confidence interval for the mean of each model suggests that one can expect to sacrifice approximately 68-90%, on average, in optimality with respect to the optimal objective value of the maximum return problem, which simply chooses the asset with the largest mean log return. We do recognise that this is an empirical investigation and that although individual cases of cost and robustness for each model are stable, particularly models R_1 to R_3 and R_8 to R_{10}, these results are dependent on one particular set of data (FTSE 100 over the period 01 January 1992 through to 01 December 2007). Furthermore, it is important to note that every model, robust or not, comes at a cost. In future work we will compare the cost of Expected value – Variance (*E-V*) portfolios to that of the robust portfolios.

References

Ben-Tal, A., Margalit, T. and Nemirovski, A., 1999. Robust Modeling of Multi-Stage Portfolio Problems. In: H. Frenk, K. Roos, T. Terlaky and S. Zhang, Eds, *High Performance Optimization.* 1st Edn. The Netherlands: Kluwer Academic Publishers, pp. 303-328.

Ben-Tal, A. and Nemirovski, A., 1998. Robust convex optimization. *Mathematics of Operations Research,* **23**(4), pp. 769-805.

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

Bertsimas, D. and Pachamanova, D., 2008. Robust multiperiod portfolio management in the presence of transaction costs. *Computers & Operations Research,* **35**(1), pp. 3-17.

Bienstock, D., 2007. Experiments in Robust Portfolio Optimization. Columbia University: http://www.columbia.edu/~dano/papers.html.

Burgman, M., 2005. Risks and Decisions for Conservation and Environmental Management. Cambridge, UK: Cambridge University Press.

Ceria, S. and Stubbs, R.A., 2006. Incorporating estimation errors into portfolio selection: Robust portfolio construction. *Journal of Asset Managment,* **7**(2), pp. 109-127.

El Ghaoui, L. and Lebret, H., 1997. Robust Solutions to Least-Squares Problems with Uncertain Data. *SIAM Journal on Matrix Analysis and Applications,* **18**(4), pp. 1035-1064.

EL Ghaoui, L., OKS, M. and Oustry, F., 2003. Worst-case Value-at-Risk and robust portfolio optimization: A conic programming approach. *Operations research,* **51**(4), pp. 543-556.

Goldfarb, D. and Iyengar, G., 2003. Robust portfolio selection problems. *Mathematics of Operations Research,* **28**(1), pp. 1.

Grant, M., Boyd, S. and Ye, Y., 2008. *CVX Users' Guide for cvx version 1.2.* Stanford University: http://www.stanford.edu/~boyd/cvx/cvx_usrguide.pdf.

Kim, S. and Boyd, S., 2007. Robust Efficient Frontier Analysis with a Separable Uncertainty Model. Stanford University: http://www.stanford.edu/~boyd/papers/pdf/rob_ef_sep.pdf.

Lobo, M. and Boyd, S., 2000. *The worst-case risk of a portfolio.* Stanford University: http://faculty.fuqua.duke.edu/%7Emlobo/bio/research-optimization.htm.

Markowitz, H., 1952. Portfolio Selection. *The Journal of Finance,* **7**(1), pp. 77-91.

Michaud, R.O., 1989. The Markowitz Optimization Enigma: Is 'Optimized' Optimal? *Financial Analysts Journal,* **45**(1), pp. 31-42.

Pflug, G. and Wozabal, D., 2007. Ambiguity in portfolio selection. *Quantitative Finance,* **7**(4), pp. 435-442.

Soyster, A. L., 1973. Convex Programming with Set-Inclusive Constraints and Applications to Inexact Linear Programming. *Operations Research,* **21**(5), pp. 1154-1157.

Tütüncü, R.H. and Koenig, M., 2004. Robust Asset Allocation. *Annals of Operations Research,* **132**(1), pp. 157-187.

Vose, D., 2000. Risk Analysis: A Quantitative Guide. 2nd Edn. West Sussex, England: John Wiley & Sons Ltd., pp. 18-27.