Minimizing CVaR and VaR for a Portfolio of Derivatives*

Siddharth Alexander**, Thomas F. Coleman***, and Yuying Li***

**Center for Applied Mathematics, Cornell University, Ithaca, New York, USA

***Department of Computer Science, Cornell University, Ithaca, New York, USA

Email: alexande@cam.cornell.edu , coleman@tc.cornell.edu, yuying@tc.cornell.edu

(This version: April 13, 2004)

Abstract. *Value at risk* (VaR) and *conditional value at risk* (CVaR) are the most frequently used risk measures in current risk management practice. As an alternative to VaR, CVaR is attractive since it is a coherent risk measure. We analyze the problem of computing the optimal VaR and CVaR portfolios. In particular, we illustrate that VaR and CVaR minimization problems for *derivatives portfolios* are typically ill-posed. For example, the VaR and CVaR minimizations based on delta-gamma approximations of the derivative values typically have an infinite number of solutions. In this paper, we focus on the portfolio selection problem which yields a portfolio of the minimum CVaR with a specified rate of return.

We propose to include cost as an additional preference criterion for the CVaR optimization problem. We demonstrate that, with the addition of a proportional cost, it is possible to compute an optimal CVaR derivative investment portfolio with significantly fewer instruments and comparable CVaR and VaR. A computational method based on a smoothing technique is proposed to solve a simulation based CVaR optimization problem efficiently. Comparison is made with the linear programming approach for solving the simulation based CVaR optimization problem.

Keywords. *derivative portfolio investment, derivative portfolio hedging, VaR, CVaR, risk minimization, Black-Scholes model, ill-posedness, transaction and management cost*

JEL Classification. *C61 and G11*

---

*This paper was presented at the International Conference on Modeling, Optimization, and Risk Management in Finance (Gainesville, March 5-7, 2003). This research was conducted using resources of the Cornell Theory Center, which is supported in part by Cornell University, New York State, and members of the Corporate Partnership Program.

1 Introduction

One of the main objectives of risk management is to evaluate and improve the performance of financial organizations in light of the risks taken to achieve profits. A current standard benchmark for firm-wide measures of risk is *value at risk* (VaR) [9]. For a given time horizon $\bar{t}$ and confidence level β, the value at risk of a portfolio is the loss in the portfolio's market value over the time horizon $\bar{t}$ that is exceeded with probability $1-\beta$. However, as a risk measure, VaR has recognized limitations. Firstly it lacks subadditivity and convexity [3, 4]. For example, the VaR of the combination of two portfolios can be greater than the sum of the VaR of the individual portfolios. Indeed, VaR is a coherent risk measure only when it is based on the standard deviation of normal distributions. In addition, it has been shown in [15, 16] that the problem of minimizing the VaR of a portfolio can have multiple local minimizers.

An alternative risk measure to VaR is *conditional value at risk* (CVaR), which is also known as *mean excess loss*, *mean shortfall* or *tail VaR*. For a given time horizon $\bar{t}$ and confidence level β, CVaR is the conditional expectation of the loss above VaR for the time horizon $\bar{t}$ and the confidence level β. The CVaR risk measure, with a slight modification, is also applicable to distributions with jumps [20]. It has been shown [18] that CVaR is a coherent risk measure that has many attractive properties including convexity, see e.g., [17] for an overview of CVaR. In addition, minimizing CVaR typically leads to a portfolio with a small VaR.

A convex optimization problem has been proposed in [19] to compute the optimal CVaR portfolio, we describe the mathematical formulation of CVaR optimization problem in §2. In particular, when this optimization problem is approximated by Monte Carlo simulation, it has an equivalent linear programming formulation and can be solved using standard linear programming methods.

Derivative contracts have become increasingly important as investment tools for achieving higher returns and decreasing funding costs. In this paper, we first analyze in §3 the well-posedness of the optimal CVaR/VaR portfolio selection problem when the investment universe consists of derivative contracts. Specifically, we illustrate that the CVaR/VaR optimization problem for derivative portfolios typically has an infinite number of solutions if the derivative values are computed using delta-gamma approximations. In particular, the optimal investment portfolios lie in a linear subspace of dimension $(n-(2d+3))$, where n is the total number of instruments in the investment universe and d is the total number of underlying risk factors; here we assume that there exist only budget and return constraints and each derivative depends on a single risk factor. Similar results are obtained if a derivative value depends on more than one risk factor; this analysis is presented in Appendix 6. Moreover, even when the derivative values are computed with more accurate methods such as analytic formulae, numerical partial differential equations, or Monte Carlo methods, the CVaR/VaR optimization problem for derivative portfolios remains ill-posed in the sense that there are many portfolios that have similar CVaR/VaR values to that of the optimal portfolio and slight perturbation of the data can lead to significantly different optimal solutions; we illustrate this with derivative CVaR portfolio examples in §3. We note that the concepts of well-posed and ill-posed problems were first introduced by Hadamard at the beginning of the 20th century [11] and ill-posed problems emerge from many areas of science and engineering, typically in the form of inverse problems.

In §4 we focus on the CVaR optimization problem and introduce cost as an additional preference; the cost is modeled as proportional to the magnitude of the holding positions. A similar consideration can be applied to the VaR optimization problem although VaR optimization is a more computationally challenging task. We show that a similar convex programming problem can be formulated for CVaR optimization under the proposed proportional cost model. In addition, we demonstrate that the proposed CVaR optimization formulation with cost is able to limit both the transaction cost and management cost; an optimal CVaR derivative investment portfolio using a suitable weighted cost parameter has smaller total trading positions, significantly fewer instruments, and comparable CVaR (and VaR).

The standard method for a CVaR optimization problem is a linear programming (LP) approach. Using Monte Carlo simulation, a piecewise linear function is used to approximate the typically continuous differentiable CVaR function which results in a linear programming problem. This LP is then solved using standard linear programming software. We illustrate that this approach becomes inefficient for large scale CVaR optimization problems.

A computational method based on a smoothing technique is proposed in §5 to efficiently solve a simulation based CVaR optimization problem. Comparison is made with the linear programming approach to solve the simulation based CVaR optimization problem. We demonstrate that the smoothing formulation, compared with the linear programming approach, is computationally much more efficient in both CPU usage and memory requirement and is capable of solving larger problems.

2 Mathematical Formulation

For a time horizon $\bar{t}$, let $f(x,S)$ denote the loss of a portfolio with decision variable $x \in \mathbb{R}^n$ and random variable $S \in \mathbb{R}^d$ denote the value of underlying risk factors at $\bar{t}$. Without loss of generality, we assume that the random variable $S \in \mathbb{R}^d$ has a probability density $p(S)$. For a given portfolio x, the probability of the loss not exceeding a threshold α is given by the cumulative distribution function

$$\Psi(x,\alpha) \stackrel{\text{def}}{=} \int_{f(x,S)\leq\alpha} p(S)dS. \tag{1}$$

When the probability distribution for the loss has no jumps, $\Psi(x,\alpha)$ is everywhere continuous with respect to α.

The VaR associated with a portfolio x, for a specified confidence level β and time horizon $\bar{t}$, is given by

$$\alpha_\beta(x) \stackrel{\text{def}}{=} \inf\{\alpha \in \mathbb{R} : \ \Psi(x,\alpha) \geq \beta\} \tag{2}$$

Note that under the assumption that $\Psi(x,\alpha)$ is everywhere continuous, there exists α (possibly not unique) such that $\Psi(x,\alpha) = \beta$.

Define $[f(x,S)-\alpha]^+$ as

$$[f(x,S)-\alpha]^+ \stackrel{\text{def}}{=} \begin{cases} f(x,S)-\alpha & \text{if } f(x,S)-\alpha > 0 \\ 0 & \text{otherwise.} \end{cases}$$

The risk measure CVaR, $\phi_\beta(x)$, is defined as [18, 20]:

$$\phi_\beta(x) \stackrel{\text{def}}{=} \inf_{\alpha} \left(\alpha + (1-\beta)^{-1}\mathbb{E}\left([f(x,S)-\alpha]^+\right)\right)$$

When the loss distribution has no jumps, CVaR is the conditional expectation of the loss, given that the loss is $\alpha_\beta(x)$ or greater, and is given by

$$\phi_\beta(x) = (1-\beta)^{-1}\int_{f(x,S)\geq\alpha_\beta(x)} f(x,S)p(S)dS. \tag{3}$$

Define the augmented function

$$F_\beta(x,\alpha) \stackrel{\text{def}}{=} \alpha + (1-\beta)^{-1}\int_{S\in\mathbb{R}^d}[f(x,S)-\alpha]^+ p(S)dS \tag{4}$$

Under the assumption that the loss function $f(\cdot,S)$ is convex, the loss distribution is continuous, it can be shown [19] that the function $F_\beta(x,\alpha)$ is convex and continuously differentiable with respect to α and $\phi_\beta(x)$ is convex with respect to x. Moreover, minimizing the CVaR over any $x \in X$, where X is a subset of $\mathbb{R}^n$, is equivalent to minimizing $F_\beta(x,\alpha)$ over $(x,\alpha) \in X \times \mathbb{R}$, i.e.,

$$\min_{x\in X}\phi_\beta(x) \equiv \min_{(x,\alpha)\in X\times\mathbb{R}} F_\beta(x,\alpha). \tag{5}$$

If, in addition, X is a convex set, then the CVaR minimization problem

$$\min_{(x,\alpha)\in X\times\mathbb{R}} F_\beta(x,\alpha) \tag{6}$$

is a convex programming problem.

3 Minimizing Risk for Derivative Portfolios

At a given time horizon $\bar{t} > 0$, assume that the underlying asset prices of the derivative instruments are $S_{\bar{t}} \in \mathbb{R}^d$, the initial asset prices are S_0, and the function $f(x,S)$ is the loss of a portfolio from a universe of n instruments. Assume that instrument values at time $\bar{t}$ are $\{V_1(S_{\bar{t}},\bar{t}),\cdots,V_n(S_{\bar{t}},\bar{t})\}$. For a portfolio selection problem and a given investment horizon $\bar{t} > 0$, the loss associated with the portfolio x is

$$f(x,S_{\bar{t}}) = -x^T\left(V^{\bar{t}} - V^0\right)$$

where for any time t, $V^t \stackrel{\text{def}}{=} [V_1(S_t,t),\ldots,V_n(S_t,t)]$. Note that $f(x,S)$ is a linear function of x and it can be easily shown that, for any $\rho > 0$,

$$\alpha_\beta(\rho\cdot x) = \rho\cdot\alpha_\beta(x), \quad \text{and} \quad \phi_\beta(\rho\cdot x) = \rho\cdot\phi_\beta(x).$$

Let $\delta V \in \mathbb{R}^n$ denote the change in the instrument values over the time horizon $\bar{t}$, i.e., $\delta V = V^{\bar{t}} - V^0$. Then the loss, $f(x,S_{\bar{t}})$, of the portfolio over the investment horizon $\bar{t}$ is $-(\delta V)^T x$.

Without loss of generality, let $x \in \mathbb{R}^n$ denote the ratio of the instrument holdings to the total initial investment wealth, i.e., x_i is the number of units of the ith instrument holding per dollar investment. (The VaR and CVaR of a portfolio with a budget ρ are simply $\rho \cdot \alpha_\beta(x)$ and $\rho \cdot \phi_\beta(x)$ respectively, where $\alpha_\beta(x)$ and $\phi_\beta(x)$ are computed for a dollar's investment).

Assume for now that the only constraints on the optimal portfolio are the budget and return constraints; the budget constraint can be expressed as

$$\left(V^0\right)^T x = 1$$

and the return constraint for the investment horizon $\bar{t}$ is

$$\left(\overline{\delta V}\right)^T x = \bar{r}$$

where $\bar{r} \geq 0$ specifies the expected return of the portfolio over the time horizon $\bar{t}$ and $\overline{\delta V} \in \mathbb{R}^n$ is the expected gain for the instruments, i.e., $\overline{\delta V} = \mathbb{E}\left[(\delta V)\right]$.

If $X = \{x : (V^0)^T x = 1, \ \left(\overline{\delta V}\right)^T x = \bar{r}\}$ is the set of feasible portfolios, we can write (6) explicitly as

$$\begin{array}{ll} \displaystyle\min_{(x,\alpha)} & \left(\alpha + (1-\beta)^{-1} \displaystyle\int_{S \in \mathbb{R}^d} [-(\delta V)^T x - \alpha]^+ p(S) dS\right) \\ \text{subject to} & \left(V^0\right)^T x = 1 \\ & \left(\overline{\delta V}\right)^T x = \bar{r}. \end{array} \tag{7}$$

We assume that a stochastic model for changes of the underlying asset prices of all the instruments in a portfolio is given. In addition, we assume that there exist methods for computing the derivative values, such as Black-Scholes formulae, delta-gamma approximations, and Monte Carlo simulation.

The continuous CVaR optimization problem (7) is a convex nonlinear minimization problem with linear constraints. If the loss distribution function is continuous, the objective function is continuously differentiable. How well is this optimization problem posed for a portfolio of derivatives?

To investigate this, let us consider the delta-gamma approximation of derivative values. For a short time horizon $t > 0$, a delta-gamma approximation can be a sufficiently accurate approximation to the derivative value and is often used in risk assessment. In general, the delta-gamma approximation describes the most significant component in the change of the derivative values and can thus provide insight into the nature of the solution. Thus we assume for now that the change, for a given horizon $\bar{t}$, in instrument values is specified by the delta-gamma approximation; for instrument i:

$$V_i^{\bar{t}} - V_i^0 = \left(\frac{\partial V_i^0}{\partial t}\right) \delta \bar{t} + \left(\frac{\partial V_i^0}{\partial S}\right)^T (\delta S) + \frac{1}{2} (\delta S)^T \Gamma_i (\delta S). \tag{8}$$

Here the vector $(\delta S) \in \mathbb{R}^d$ denotes the change in the underlying values, $\frac{\partial V_i^0}{\partial t}$ denotes the initial theta sensitivity of the ith instrument value to time, $\frac{\partial V_i^0}{\partial S} \in \mathbb{R}^d$ denotes the initial delta sensitivity of the ith instrument with respect to the underlyings, and $\Gamma_i \in \mathbb{R}^{d \times d}$ is the

Hessian matrix denoting the initial gamma sensitivity of the ith instrument with respect to the underlyings, and δt is change in time.

Let $\frac{\partial V^0}{\partial t}$ and $\frac{\partial V^0}{\partial S}$ denote the initial sensitivities for all instruments in the investment universe:

$$\begin{aligned} \frac{\partial V^0}{\partial t} &\stackrel{\text{def}}{=} \left[\frac{\partial V_1^0}{\partial t}, \cdots, \frac{\partial V_n^0}{\partial t}\right] \in \mathbb{R}^n \\ \frac{\partial V^0}{\partial S} &\stackrel{\text{def}}{=} \left[\frac{\partial V_1^0}{\partial S}, \cdots, \frac{\partial V_n^0}{\partial S}\right]^T \in \mathbb{R}^{n\times d} \end{aligned}$$

Assume for now that each instrument depends on a single risky asset. If a derivative value depends on more than one risk factor, similar results can be obtained by accounting for the cross-partial derivatives; this analysis is presented in Appendix A. In the case of a single risk factor, the only non-zero entries in the vector $\frac{\partial V_i^0}{\partial S}$ and matrix Γ_i are entries i and (i,i) respectively. Let

$$\Gamma \stackrel{\text{def}}{=} \left[\Gamma_1^{diag}, \cdots, \Gamma_n^{diag}\right]^T \in \mathbb{R}^{n\times d}$$

where Γ_i^{diag} represents the diagonal of the matrix Γ_i as a column vector. Let $(\delta S)^2$ be the vector with each entry of δS squared. If we set

$$\Lambda \stackrel{\text{def}}{=} \left[\left(\frac{\partial V^0}{\partial t}\right), \left(\frac{\partial V^0}{\partial S}\right), \frac{1}{2}\Gamma\right] \in \mathbb{R}^{n\times(2d+1)}, \tag{9}$$

the loss in portfolio value is given by

$$f(x,S) = -x^T\Lambda \begin{bmatrix} \delta \bar{t} \\ \delta S \\ (\delta S)^2 \end{bmatrix} \tag{10}$$

If $n > 2d+1$, there exists a non-zero $z \in \mathbb{R}^n$ satisfying $\Lambda^T z = 0$. It is clear that, for any θ,

$$f(x,S) = f(x+\theta\cdot z, S), \quad \forall S.$$

Thus the portfolios x and $(x+\theta\cdot z)$ have the same VaR and CVaR under the delta-gamma approximation. For a portfolio selection problem, if $X = \{x : (V^0)^T x = 1, \ \left(\overline{\delta V}\right)^T x = \bar{r}\}$ denotes the set of feasible portfolios corresponding to the budget and return constraints, we may deduce that if $n > (2d+3)$, then the optimal CVaR and VaR portfolios for the selection problem defined by $\min_{x\in X}\phi_\beta(x)$ and $\min_{x\in X}\alpha_\beta(x)$, for any $0 < \beta < 1$, lie in a linear subspace of dimension $n-(2d+3)$. This implies that the VaR and CVaR derivative minimization problems, under these stated assumptions, are ill-posed and different computational methods may produce different optimal portfolios.

We note that not all derivative values can be expressed as simply $V(S_t,t)$ where S_t represents risk factor values at time t. Asian options for example have a strong dependency on the history of the stock price. The analysis when such instruments are present is more complex but similar results may be obtained.

When the derivative values are computed through more accurate methods, such as analytic formulae or Monte Carlo simulation, the CVaR optimization problem typically remains ill-posed in the sense that there are many portfolios that have similar risk values to that of the optimal portfolio and slight perturbation of the data can lead to significantly different optimal solutions. We will subsequently illustrate this with an example.

Although much of the subsequent discussion is applicable to the VaR optimization problem, we will focus only on the CVaR optimization problem due to its computational tractability.

A continuous CVaR optimization problem (7) can be approximated using Monte Carlo simulation. Assume that $\{(\delta V)_i\}_{i=1}^m$ are independent samples of δV, the change in the instrument values over the given horizon. Then the following is an approximation to the optimization problem (7):

$$\min_{(x,\alpha)\in X\times\mathbb{R}} \left(\bar{F}_\beta(x,\alpha) \stackrel{\text{def}}{=} \alpha + \frac{1}{m(1-\beta)} \sum_{i=1}^{m} [-(\delta V)_i^T x - \alpha]^+ \right). \tag{11}$$

When the subset X is specified by a finite set of linear constraints, (11) has an equivalent linear programming formulation which can be solved by standard methods for linear programming, see e.g., [19] and [14] for an overview of modeling the CVaR problem (as well as other risk measures) as an LP. Here we use interior point method software MOSEK [1].

Naturally, additional properties can be included in the CVaR optimization problem as constraints to alleviate the ill-posedness of the problem and produce a more desirable optimal portfolio. However, one needs to be careful to ensure that these constraints are meaningful and consistent in the sense that there exist feasible solutions. In addition, simply adding constraints may give a false sense of security; the optimization problem may remain ill-posed, as will be illustrated next.

The most natural constraints that one can add are simple bound constraints on the instrument holdings. The following example illustrates that this does not necessarily regularize the ill-posedness of the problem. In addition, it demonstrates significant consequences of the ill-posedness of the CVaR derivative portfolio optimization problem.

Assume, for example, that the feasible portfolios correspond to ones that satisfy budget and return constraints as well as bound constraints on the instrument positions. Then the optimization problem (11) becomes

$$\begin{array}{ll} & \min\limits_{(x,\alpha)} \left(\alpha + \dfrac{1}{m(1-\beta)} \displaystyle\sum_{i=1}^{m} [-(\delta V)_i^T x - \alpha]^+ \right) \\ \text{subject to} & \left(V^0\right)^T x = 1 \\ & \left(\overline{\delta V}\right)^T x = \bar{r} \\ & l \le x \le u. \end{array} \tag{12}$$

First we demonstrate how CVaR optimal portfolios change with slight perturbations to the initial volatility. Assume that the initial underlying asset price is $S_0 = 100$ and is log-normally distributed, the volatility is 0.2, and the drift is 0.1. The annual risk free interest rate is 5%, a trading year consists of 252 days, and our time horizon is 10 days. We select our portfolio from 15 different call options on the underlying asset priced using Black-Scholes type formulae. The options are described by all combinations of strike prices ($[0.9; 0.95; 1; 1.05; 1.1] \times S_0$) and expiries ($[30; 50; 70]$ days). We do not allow long or short

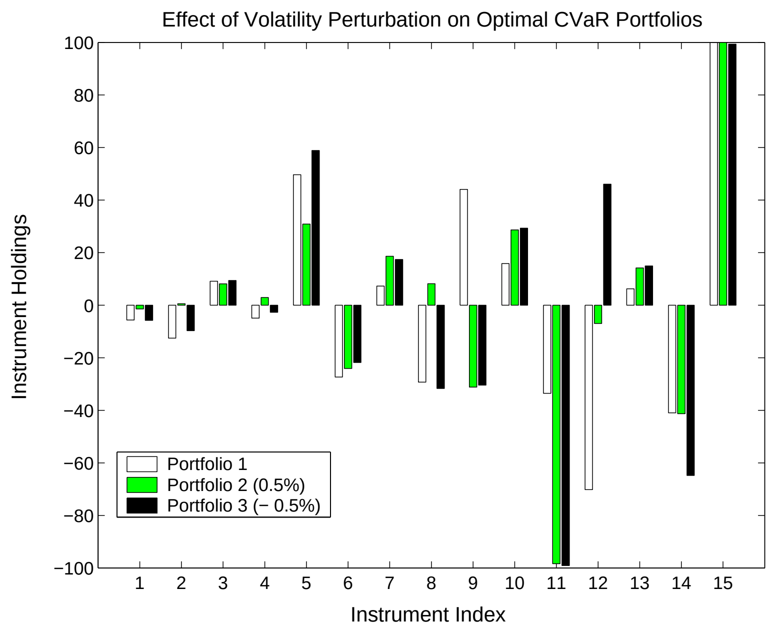

Figure 1: Optimal Portfolio Holdings are Sensitive to Volatility Perturbation: $\beta = 0.95$, $\bar{t} = 10$ days

positions of more than 100 units on any option. Our return constraint is twice the risk-free return over the investment horizon and we use a $\beta = 0.95$ confidence level. For this example we use $m = 25000$ Monte Carlo samples. For a \$1 investment, the optimal portfolio under these assumptions has VaR=0.00412 and CVaR=0.00421. Figure 1 illustrates the optimal portfolio positions for different volatility perturbations; portfolio 1 corresponds to no perturbation and portfolios 2 and 3 correspond to perturbations of 0.5% and -0.5% respectively. In this case, the new VaR and CVaR values are VaR=0.00416, CVaR=0.00423 and VaR=0.00407, CVaR=0.00414 respectively. Note that while the risk values have changed by less than 2%, the portfolios are significantly different.

To further illustrate the properties of the optimal portfolio from (12), we consider a universe of 196 instruments consisting of 12 vanilla calls, 12 vanilla puts, 12 binary calls, and 12 binary puts on each of four correlated assets, and the four underlying assets themselves; here the derivative instruments are all European options. The initial asset prices, the covariance matrix of the annual returns, and the expected rates of return of the four assets are given in Tables (9)-(11) in Appendix B respectively. In this paper, we use these specifications for all our computational results, varying only the option types being considered. The derivatives are priced using Black-Scholes type formulae, assuming that the underlying prices are log-normally distributed. For this example, we use $m = 25000$ Monte Carlo samples. The strike prices used for options on each asset are $[0.8; 1.025; 1.25] \times S_0$ where S_0 is the time 0 asset values (see Table 9 in Appendix B). The times to expiry are $[2; 4; 6; 8] \times \bar{t}$, where $\bar{t} = 10$ days is the investment horizon (we assume that there are 250 trading days in a year). The options are all combinations of strikes and expiry times. The required portfolio return is twice the risk free interest rate over the investment horizon with the annual risk free interest

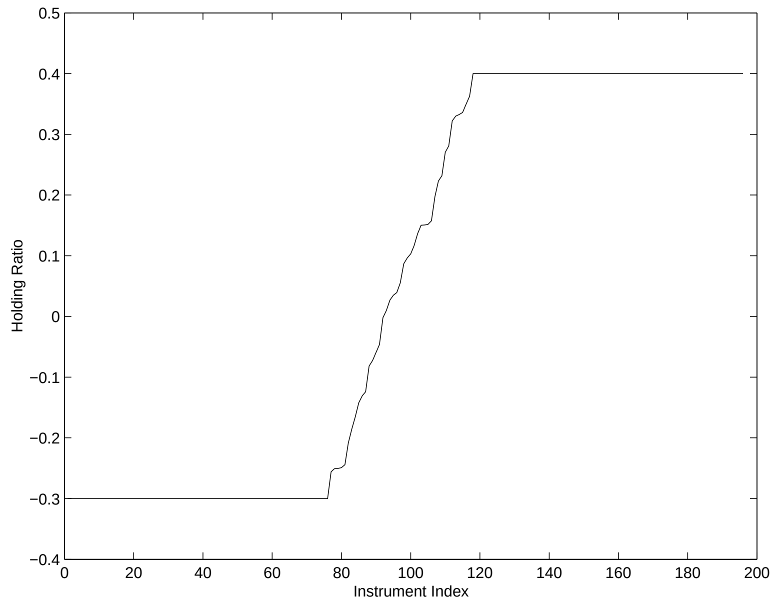

Figure 2: (Ordered) Holding Ratios From (12) With No Cost Consideration: $\beta = 0.95$, $\bar{t} = 10$ days

rate equal to 5%. We use lower bounds of -0.3 and upper bounds of 0.4 for this example. For an investment of \$100, no more than 30 units of each instrument can be shorted and no more than 40 units of each instrument can be bought. An (ordered) optimal portfolio holding ratio x^* from (12), computed using the interior point software MOSEK, is shown in Figure 2 [1]. Note that MOSEK uses an interior-point optimizer to compute a solution and hence not all instruments are at their bounds. The optimal portfolios computed by CPLEX have more instruments at their bounds, but are numerically similar to the MOSEK optimal portfolios.

Let us exclude the ith instrument from the optimal portfolio if $|x_i^*| \leq 10^{-5}$. We first observe that the optimal portfolio consists of all the instruments in the investment universe. In addition, about 77% of instrument holding ratios are equal to either their upper or lower limits. Such an optimal portfolio is undesirable in that it leads to large transaction as well as management costs. Moreover, any model error will be magnified for a portfolio with extreme holdings; this is illustrated for the portfolio hedging problem in [2]. Indeed, the optimization problem with the bound constraints remains ill-posed in the sense that there are many different portfolios with similar CVaR values; we will illustrate this next by providing a more desirable optimal CVaR formulation which produces more attractive portfolios with similar CVaR values.

In addition to the portfolio selection problem, another practically important derivative portfolio optimization problem is the portfolio hedging problem. In this context, for a given hedging horizon $\bar{t}$, one has an initial portfolio and an associated portfolio loss $\Pi^0(S, \bar{t})$. The goal is to decrease the risk of this portfolio by selecting an appropriate hedging portfolio from

[1] The instruments are ordered so as to illustrate the number of holding ratios that are significantly larger than zero, in particular, the number of holding ratios that are at their bounds.

the available instruments $\{V_1(S,t), \cdots, V_n(S,t)\}$. Thus the loss function for the hedging problem has the form

$$f(x,S) = \Pi^0(S,\bar{t}) - x^T\left(V(S,\bar{t}) - V(S_0,0)\right).$$

While we devote our analysis of ill-posedness in this paper to the portfolio selection problem, similar analysis applies to the portfolio hedging problem where the goal is to hedge a given portfolio using more liquid derivatives with CVaR as the risk measure; in [2], without the mathematical analysis in the general setting and computational results for the smoothing technique, a shorter and simpler paper is written to illustrate the effect of the ill-posedness in the derivative portfolio hedging problem based on CVaR; hedging performances of the optimal portfolios under different cost considerations are also compared. For the hedging computational results in [2], a shorter time horizon is considered and no return constraint or budget constraint is imposed.

4 CVaR Optimization with Cost

Given that the CVaR optimization problem for a portfolio of derivatives is ill-posed, in order to generate a stable solution, additional meaningful criteria need to be considered for a derivative portfolio CVaR optimization problem. A natural meaningful consideration in portfolio investment or risk management is transaction and management cost. A portfolio, which, in addition to a small CVaR, incurs a small transaction and management cost, is certainly more attractive. We can regard the management cost as a function of the number of (non-zero holding) instruments in a portfolio. Unfortunately, it is difficult to include this explicitly into an optimization formulation since it is computationally challenging to solve the resulting mixed integer program. Our objective is to seek a portfolio which consists of a small number of instruments by minimizing a combination of CVaR and a suitable cost function without the need to solve a mixed integer programming problem.

Let us assume that the cost of holding an instrument is proportional to the magnitude of the instrument holdings. Then we seek a portfolio which has a minimum weighted combination of CVaR and the proportional cost:

$$\min_{x\in X} \left(\phi_\beta(x) + \sum_{i=1}^{n} c_i|x_i|\right) \tag{13}$$

where $\phi_\beta(x)$ is as defined in (3). Here $c \geq 0$ is the weighted cost, representing the cost as well as the tradeoff between minimizing CVaR and cost.

The weighted cost parameter $c_i \geq 0$ can be interpreted as a measure of relative desirability to exclude the ith instrument from the optimal portfolio: if c_i is greater than some finite threshold value, and there exists a feasible portfolio with $x_i = 0$, then the optimal portfolio x^* for (13) is guaranteed to exclude the ith instrument, i.e. $x_i^* = 0$. In this sense, we can regard our cost model as a model for management cost. This property of the cost model (13) is due to the fact that the objective function $(\phi_\beta(x) + \sum_{i=1}^n c_i|x_i|)$ is an exact penalty function of a constrained optimization problem. We refer interested readers to [10] for a more detailed discussion on the exact penalty function. Note that if one models the cost as

$\sum_{i=1}^n c_i x_i^2$ for example, the resulting optimal portfolio typically has few, if any at all, of its instruments with a small holding ratio $|x_i^*|$ (e.g., $|x_i^*| \leq 10^{-5}$). For the quadratic penalty function, the constraint $x_i^* = 0$ is only satisfied as the penalty parameter c_i tends to $+\infty$.

To solve (13), we can similarly consider the augmented function $F_\beta(x, \alpha) + \sum_{i=1}^n c_i|x_i|$. It is clear that $F_\beta(x, \alpha) + \sum_{i=1}^n c_i|x_i|$ remains convex and continuously differentiable with respect to α since $\sum_{i=1}^n c_i|x_i|$ is convex and has no dependence on α; the analysis of [19] applies. Moreover minimizing the sum of the weighted cost and CVaR of a portfolio x in any subset X of $\mathbb{R}^n$ is equivalent to minimizing $F_\beta(x, \alpha) + \sum_{i=1}^n c_i|x_i|$ over $(x, \alpha) \in X \times \mathbb{R}$, i.e.,

$$\min_{x \in X} \left(\phi_\beta(x) + \sum_{i=1}^n c_i|x_i| \right) \equiv \min_{(x,\alpha) \in X \times \mathbb{R}} \left(F_\beta(x, \alpha) + \sum_{i=1}^n c_i|x_i| \right).$$

In addition, $F_\beta(x, \alpha) + \sum_{i=1}^n c_i|x_i|$ is convex with respect to (x, α) and $\phi_\beta(x) + \sum_{i=1}^n c_i|x_i|$ is convex with respect to x if the loss function $f(x, S)$ is convex with respect to x. Moreover, if X is a convex set, the minimization problem

$$\min_{(x,\alpha) \in X \times \mathbb{R}} \left(F_\beta(x, \alpha) + \sum_{i=1}^n c_i|x_i| \right) \tag{14}$$

is a convex programming problem.

When (14) is approximated through Monte Carlo simulation, and X is specified by the budget and return constraints and bounds on the holding ratios x, the CVaR optimization problem with a proportional cost becomes a constrained piecewise linear minimization problem:

$$\begin{array}{ll} & \min\limits_{(x,\alpha)} \left(\alpha + \dfrac{1}{m(1-\beta)} \sum\limits_{i=1}^m [-(\delta V_i)^T x - \alpha]^+ + \sum\limits_{j=1}^n c_j|x_j| \right) \\ \text{subject to} & \left(V^0\right)^T x = 1 \\ & \left(\overline{\delta V}\right)^T x = \bar{r} \\ & l \leq x \leq u \end{array} \tag{15}$$

To illustrate the effect of the weighted cost parameter c on the optimal portfolio obtained from the CVaR cost model (14), we consider the weighted cost parameter $c_i = \omega \cdot |\overline{\text{CVaR}}^0|$, $1 \leq i \leq n$, where $\overline{\text{CVaR}}^0$ denotes the optimal CVaR from (15) with no cost consideration, for a dollar's investment. (Here we are implicitly assuming that the transaction costs of instruments are the same).

Consider the same 196 instrument example in §3. We first recall that the optimal CVaR portfolio, under no cost consideration, contains all the 196 instruments. In addition 77% of holding ratios are at their bounds. Figure 3 plots the optimal portfolio holding ratio x^*, for the same example, for the weighted cost $c_i = \omega \cdot |\overline{\text{CVaR}}^0|$ where $\omega = 0,\ 0.005,$ and 0.01. We note that for $\omega = 0.005$ and 0.01, the optimal portfolios are preferable in the sense that they contain only 35.7% and 29.1% of the 196 instruments respectively.

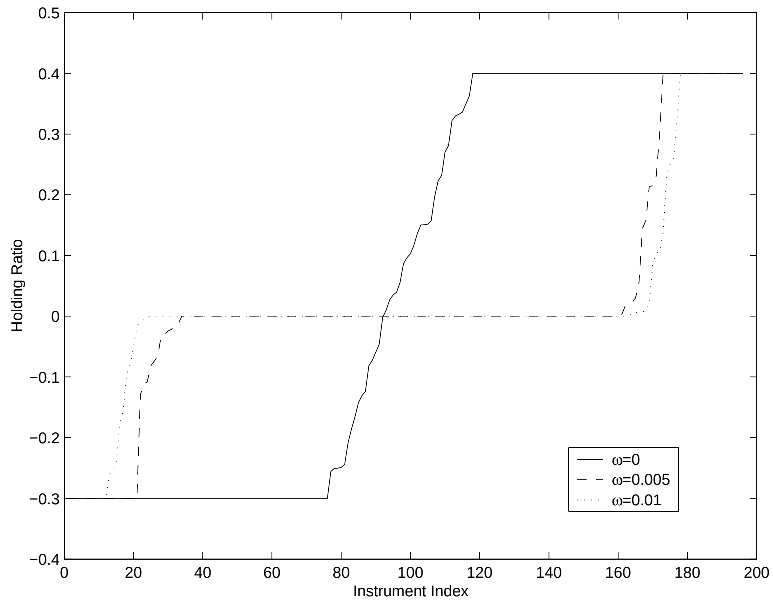

Figure 3: Holding Ratios With Varying Costs: $\beta = 0.95$, $\bar{t} = 10$ days

In order to analyze the impact of the cost consideration on risks, we consider the relative differences of VaR and CVaR under different weighted cost parameters with respect to that under cost consideration, i.e.,

$$\text{RelDifVaR}(\omega) \stackrel{\text{def}}{=} \left| \frac{\text{VaR}(\omega) - \text{VaR}^0}{\text{VaR}^0} \right|, \tag{16}$$

$$\text{RelDifCVaR}(\omega) \stackrel{\text{def}}{=} \left| \frac{\text{CVaR}(\omega) - \text{CVaR}^0}{\text{CVaR}^0} \right|, \tag{17}$$

where CVaR^0 is the optimal CVaR value from (15) under no cost consideration. Throughout the paper, the VaR and CVaR reported correspond to the VaR and CVaR for the Monte Carlo approximation. Note that the loss distribution from simulations has jumps. Consider the loss associated with a portfolio x for m scenarios, $(\text{loss})_1 \leq \ldots \leq (\text{loss})_m$, with each loss $(\text{loss})_i$ having probability p_i. For a confidence level β, let $i_\beta \leq m$ be the index such that

$$\sum_{i=1}^{i_\beta} p_i \geq \beta > \sum_{i=1}^{i_\beta - 1} p_i. \tag{18}$$

Then the VaR is given by $\alpha_\beta(x) = (\text{loss})_{i_\beta}$ and CVaR equals

$$\phi_\beta(x) = \frac{1}{1-\beta} \left[\left(\sum_{i=1}^{i_\beta} p_i - \beta \right) \alpha_\beta(x) + \sum_{i=i_\beta+1}^{m} p_i (\text{loss})_i \right]. \tag{19}$$

For a more detailed discussion of CVaR for scenario models, see [20].

β	ω	VaR	CVaR	RelDifVaR	RelDifCVaR	# instr
	0.000	1.5795	1.6458	0.0000	0.0000	196
	0.005	1.6544	1.7251	0.0474	0.0482	70
0.95	0.010	1.6990	1.7701	0.0757	0.0755	57
	0.050	1.9933	2.0545	0.2620	0.2483	34
	0.100	2.3124	2.3640	0.4641	0.4364	26
	0.000	1.6891	1.7171	0.0000	0.0000	196
	0.005	1.7654	1.7910	0.0452	0.0431	72
0.99	0.010	1.8066	1.8328	0.0696	0.0674	61
	0.050	2.1271	2.1443	0.2593	0.2488	34
	0.100	2.4067	2.4172	0.4249	0.4077	22

Table 1: Effect of Weighted Cost Parameters on the Optimal CVaR Portfolio for $\bar{t} = 10$ days

β	ω	VaR	CVaR	RelDifVaR	RelDifCVaR	# instr
	0.000	3.1164	3.1626	0.0000	0.0000	196
	0.005	3.2449	3.2934	0.0412	0.0414	90
0.95	0.010	3.3414	3.3950	0.0722	0.0735	76
	0.050	4.1099	4.1456	0.3188	0.3108	43
	0.100	4.8562	4.8682	0.5583	0.5393	20
	0.000	3.1852	3.2134	0.0000	0.0000	196
	0.005	3.3124	3.3408	0.0399	0.0396	90
0.99	0.010	3.4115	3.4379	0.0710	0.0699	77
	0.050	4.2505	4.2739	0.3344	0.3300	37
	0.100	4.8646	4.8854	0.5272	0.5203	19

Table 2: Effect of Weighted Cost Parameters on the Optimal CVaR Portfolio for $\bar{t} = 62.5$ days

Tables 1-3 tabulate relative risk differences for $\beta = 0.95$ and $\beta = 0.99$ with different weighted cost parameters for an investment horizon of 10, 62.5, and 125 days respectively. The investment universe consists of 196 instruments which have different expiries depending on the investment horizon, see Appendix B for detailed specifications. The VaR and CVaR reported here are for investment portfolios with an initial wealth of \$100. The results are for a single simulation problem and do not represent averages. Tables 1-3 illustrate that, using the CVaR and cost optimization formulation (15), it is possible to obtain CVaR optimal portfolios with significantly fewer instruments but comparable risks. For example, for $\bar{t} = 10$ days with $\omega = 0.005$, the optimal risks reflect an increase of less than 5% compared to that under no cost. Given the inevitable existence of model error as well as computational error due to, e.g., Monte Carlo approximation, these small differences in risk may be entirely acceptable. The number of instruments in the optimal portfolio, however, is less than 36% of the number of non-zero holdings under no cost, assuming a cutoff of 10^{-5}. Not surprisingly, we observe that, as the cost parameter increases, the risk increases and the number of instruments in the optimal portfolio decreases.

β	ω	VaR	CVaR	RelDifVaR	RelDifCVaR	# instr
0.95	0.000	3.6263	3.6817	0.0000	0.0000	196
	0.005	3.7915	3.8606	0.0456	0.0486	76
	0.010	3.9325	4.0013	0.0844	0.0868	61
	0.050	4.6119	4.6379	0.2718	0.2597	30
	0.100	4.8658	4.8793	0.3418	0.3253	24
0.99	0.000	3.7094	3.7432	0.0000	0.0000	196
	0.005	3.8420	3.8715	0.0357	0.0343	87
	0.010	4.0259	4.0641	0.0853	0.0857	65
	0.050	4.7118	4.7361	0.2702	0.2653	28
	0.100	4.9441	4.9652	0.3329	0.3265	17

Table 3: Effect of Weighted Cost Parameters on the Optimal CVaR Portfolio for $\bar{t} = 125$ days

Although we have included a cost consideration as a property of the optimal portfolio in the objective function in (15), we have excluded the actual transaction cost from the budget constraint in our optimal CVaR formulation since the actual transaction cost typically constitutes a very small portion of the budget. In addition, we caution that one needs to be careful when adding the transaction cost into the formulation; otherwise it may lead to a nonconvex programming problem. To illustrate this potential difficulty, assume that we include the transaction cost only in the budget constraint. We note that, alternative to (7), two equivalent formulations for the CVaR optimization problem are

$$\begin{array}{ll} & \min_{(x,\alpha)\in X\times\mathbb{R}} \quad \hat{\phi}\cdot F_\beta(x,\alpha) - \left(\overline{\delta V}\right)^T x \\ \text{subject to} & \left(V^0\right)^T x = 1 \end{array} \tag{20}$$

and maximizing return subject to a constraint on CVaR (21)

$$\begin{array}{ll} & \max_{(x,\alpha)\in X\times\mathbb{R}} \quad \left(\overline{\delta V}\right)^T x \\ \text{subject to} & \left(V^0\right)^T x = 1 \\ & F_\beta(x,\alpha) \le \bar{\phi}. \end{array} \tag{21}$$

Here $\bar{\phi}$ and $\hat{\phi}$ are two positive constants.

Consider the formulation (22) of maximizing return subject to a bound on CVaR and a budget constraint that accounts for transaction costs:

$$
\begin{aligned}
& \min_{(x,\alpha)} \quad -\left(\overline{\delta V}\right)^T x \\
\text{subject to} \quad & \bar{F}_\beta(x,\alpha) \le \bar{\phi} \\
& \sum_{i=1}^n \tilde{c}_i V_i^0 |x_i| + \sum_{i=1}^n V_i^0 x_i = 1 \\
& l \le x \le u
\end{aligned} \tag{22}
$$

where $\tilde{c}_i$ is the actual transaction cost per unit holding ratio for the ith instrument, a linear transaction cost structure. We refer readers to [6] and [8] for a more detailed analysis of portfolio optimization with transaction costs and [13] for a non-convex transaction cost structure. Here $\bar{F}_\beta(x,\alpha)$ is as defined in (11). For simplicity of illustration, let us consider the problem of computing the portfolio with the maximum return, i.e., $\bar{\phi} = \infty$. Then (22) simplifies to

$$
\begin{aligned}
& \min_{x} \quad -\left(\delta V\right)^T x \\
\text{subject to} \quad & \sum_{i=1}^n \tilde{c}_i V_i^0 |x_i| + \sum_{i=1}^n V_i^0 x_i = 1 \\
& l \le x \le u
\end{aligned} \tag{23}
$$

Eliminating the absolute value in the budget constraint, an equivalent formulation to (23) is :

$$
\begin{aligned}
& \min_{x} \quad -\left(\overline{\delta V}\right)^T x \\
\text{subject to} \quad & \sum_{i=1}^n \tilde{c}_i V_i^0 (x_i^+ + x_i^-) + \sum_{i=1}^n V_i^0 x_i = 1 \\
& x = x^+ - x^- \\
& x_i^+ x_i^- = 0 \qquad i = 1, \ldots, n \\
& x^+ \ge 0, x^- \ge 0 \\
& l \le x \le u
\end{aligned} \tag{24}
$$

Note that the optimization problem (24) now becomes nonconvex due to the presence of the constraints $x_i^+ x_i^- = 0,\ i = 1, \ldots, n$. One may be attempted to simply discard the constraints $x_i^+ x_i^- = 0,\ i = 1, \ldots, n$, given that it is never optimal to buy and sell the same instrument simultaneously. This then leads to solving

$$
\begin{array}{lll}
& \min_{x,x^+,x^-} & -\left(\overline{\delta V}\right)^T x \\
\text{subject to} & & \sum_{i=1}^{n} \tilde{c}_i V_i^0 (x_i^+ + x_i^-) + \sum_{i=1}^{n} V_i^0 x_i = 1 \\
& & x = x^+ - x^- \\
& & x^+ \geq 0, x^- \geq 0 \\
& & l \leq x \leq u
\end{array}
\tag{25}
$$

We note that, for the formulation (22) in which the transaction cost only appears in the budget constraint, solving (25) is not equivalent to solving (22). The reason is that, a solution to (25) may not satisfy the quadratic constraints $x_i^+ x_i^- = 0$ and is thus not necessarily a solution to (24). This is illustrated by the following simple counter-example with two instruments. Let $\overline{\delta V} = [0.1; -0.2]$, $V^0 = [1; 1]$, $\tilde{c} = [0.01; 0.01]$, $l = -[0.5; 0.5]$, and $u = [1; 1]$. Then $x = [1; -0.5]$ is a solution to (25)(with $x^+ = [15.1874; 10.0626]$ and $x^- = [14.1874; 10.5626]$). Note, however, that $x = [1; -0.5]$ is not a solution to (23) since it does not satisfy the budget constraint. The minimizer of (23) can be shown to be $x = [1; -\frac{1}{99}]$. If one adds the transaction cost in the budget constraint, one may need to include the transaction cost in the objective function in (22) to ensure that the problem can still be solved as a linear programming problem.

5 Minimizing CVaR Efficiently

The simulation CVaR optimization problem (15) is a piecewise linear minimization problem subject to linear constraints. As discussed previously, one way of computing a solution to (15) is to solve an equivalent linear programming problem:

$$
\begin{array}{lll}
& \min_{(x,y,z,\alpha)} & \left(\alpha + \frac{1}{m(1-\beta)} \sum_{i=1}^{m} y_i + \sum_{j=1}^{n} c_j z_j \right) \\
\text{subject to} & & \left(V^0\right)^T x = 1 \\
& & (\delta V)^T x = \bar{r} \\
& & y \geq -Bx - \alpha e_m, \\
& & z - x \geq 0, \;\; z + x \geq 0 \\
& & l \leq x \leq u, \;\; y \geq 0
\end{array}
\tag{26}
$$

where the m-by-n scenario loss matrix B is given by

$$
B = [(\delta V)_1^T; (\delta V)_2^T; \cdots; (\delta V)_m^T]
$$

and $e_m \in \mathbb{R}^m$ is the vector of all ones. This linear program has $O(m+n)$ variables and $O(m+n)$ constraints, where m is the number of Monte Carlo samples and n is the number of instruments. We assume that the loss (δV) is computed using computational methods such as analytic formulae and Monte Carlo simulation.

Linear programming is the simplest constrained optimization problem; there exists, for this class of problems, the most thorough theoretic analysis and reliable software. There are two main types of methods to solve a linear programming problem, simplex methods [7] and interior point methods [12]. The following are the most relevant properties of these two methods to our CVaR optimization problem. Consider a standard linear program $min\{c^T z \mid Az = b, z \geq 0\}$, where A is an $M \times N$ full rank matrix. Simplex methods compute a solution in a finite number of iterations by following a path from vertex to vertex along the edges of the polyhedron representing the feasible region. For a linear program of N variables, each iteration of a simplex method performs $O(N^2)$ computation and typically the method requires a large number of iterations (roughly between $2M$ to $3M$). An interior point method, on the other hand, produces an infinite sequence of approximations which converge to a solution in the limit. Interior point methods are shown to have polynomial complexity. They require $O(N^3)$ computation per iteration and the number of iterations can be bounded by $O(\sqrt{N}L)$ where L is the input length for integer data. For the CVaR portfolio optimization problem, a potential advantage of the simplex method is its ability to use a warm start. Generating a starting point, on the other hand, is an important part of an interior point method. It is not clear how a standard interior point method would utilize a warm start, if at all possible.

Although it is known that both CPLEX and MOSEK are capable of solving very large linear programming problems in a short amount of time, the efficiency of both methods depends heavily on the *sparsity structures* of the problem. The linear programming problem arising from the CVaR optimization problem has a large dense block; the size of this dense block is determined by the number of scenarios and the number of instruments. We illustrate below that the computational cost for solving a CVaR problem via the linear programming approach quickly becomes prohibitive as the number of simulations and/or instruments become large.

Table 4 illustrates how the cpu time grows with the number of simulations and the number of instruments for the CVaR optimization problem (12). The comparison is made between CPLEX version 6.6 which implements a simplex method and the MOSEK Optimization Toolbox for MATLAB version 6 (for Solaris Sparc) which implements an interior point method for single problems. The problems are implemented in MATLAB version 6.1 and run on a Sun Sparc Ultra-2 machine.

Table 4 clearly illustrates that, using the standard linear programming software, the computational cost as well as the memory requirement quickly become prohibitive as the number of Monte Carlo samples and the number of instruments increase. For example, with 200 instruments and more than $25,000$ simulations, a significant amount of the elapsed time is spent in swapping relevant data in and out of the cache memory. With 200 instruments and $50,000$ scenarios, the elapsed time is significantly longer than that of the 48 instrument example, with the memory swapping dominating the elapsed time, and the entry is marked by " - " in the table.

As an alternative to the linear programming approach for the CVaR optimization problem, we investigate a computationally efficient method which directly exploits the property of the CVaR optimization problem; our ultimate objective is to be able to solve large scale CVaR

	MOSEK (cpu sec)			CPLEX (cpu sec)		
	# of instruments being considered					
# scenarios	8	48	200	8	48	200
10000	11.07	61.96	1843.90	53.68	427.97	2120.84
25000	30.02	162.13	14744.64	351.44	2345.43	9907.99
50000	43.62	642.24	-	1673.82	9296.98	-

Table 4: CPU time for standard LP methods: $\beta = 0.99$

portfolio problems.

We want to solve a derivative portfolio CVaR optimization problem

$$\min_{(x,\alpha)\in X\times\mathbb{R}} \left(F_\beta(x,\alpha) + \sum_{i=1}^{n} c_i |x_i| \right).$$

through Monte Carlo simulation. We assume subsequently that the cumulative loss distribution function is continuous. The augmented CVaR function $F_\beta(x,\alpha) + \sum_{i=1}^{n} c_i|x_i|$ is continuously differentiable under the assumption that the loss distribution has no jumps. The linear programming approach arises from approximating the continuously differentiable function $F_\beta(x,\alpha)$ by the piecewise linear objective function

$$\bar{F}_\beta(x,\alpha) = \alpha + \frac{1}{m(1-\beta)} \sum_{i=1}^{m} [-(\delta V)_i^T x - \alpha]^+.$$

As the number of Monte Carlo simulations increases, the piecewise linear approximation $\bar{F}_\beta(x,\alpha)$ approaches the continuously differentiable function $F_\beta(x,\alpha)$.

As an alternative to the piecewise linear approximation $\bar{F}(x,\alpha)$, we consider a continuously differentiable piecewise quadratic approximation $\tilde{F}(x,\alpha)$ to the continuously differentiable function $F(x,\alpha)$. Let

$$\tilde{F}_\beta(x,\alpha) \stackrel{\text{def}}{=} \alpha + \frac{1}{m(1-\beta)} \sum_{i=1}^{m} \rho_\epsilon(-(\delta V)_i^T x - \alpha) \tag{27}$$

where $\rho_\epsilon(z)$ is a continuously differentiable piecewise quadratic function which approximates the piecewise linear function $\max(z,0)$: given a resolution parameter $\epsilon > 0$,

$$\rho_\epsilon(z) \stackrel{\text{def}}{=} \begin{cases} z & \text{if } z \geq \epsilon \\ \frac{z^2}{4\epsilon} + \frac{1}{2}z + \frac{1}{4}\epsilon & \text{if } -\epsilon \leq z \leq \epsilon \\ 0 & \text{otherwise.} \end{cases} \tag{28}$$

To illustrate the smoothness of $\bar{F}(x,\alpha)$ and $\tilde{F}(x,\alpha)$, let us consider the function $g(\alpha) = \mathbf{E}([S-\alpha]^+)$ assuming that S is a standard normal. Figure 4 graphically illustrates the accuracy and smoothness of the approximations

$$\frac{1}{m}\sum_{i=1}^{m} [S_i - \alpha]^+$$

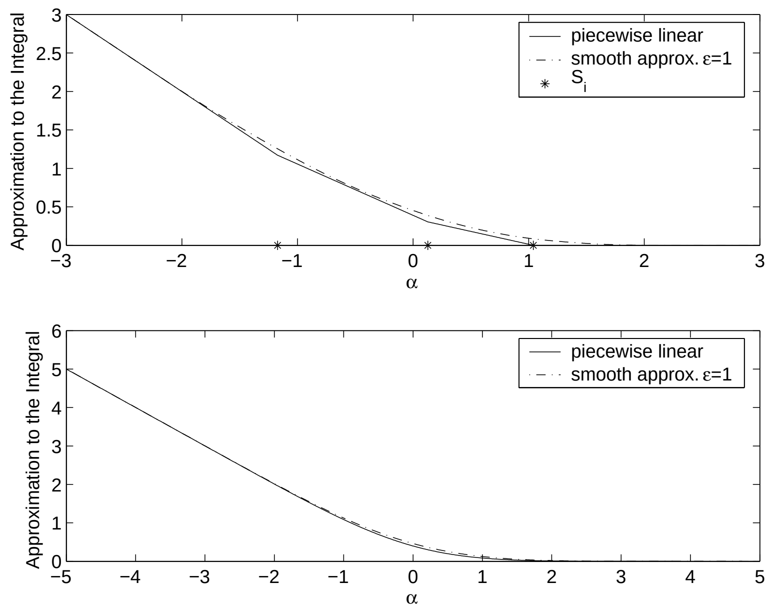

Figure 4: A Smooth Approximation

and $\frac{1}{m}\sum_{i=1}^{m}\rho_\epsilon(S_i-\alpha)$ as compared to $g(\alpha)$; the top subplot is for $m=3$ (the asterisks on the x-axis represent the S_i) and the bottom subplot is for $m=10,000$. It can be observed that, as the number of independent samples m increases, the difference between $\frac{1}{m}\sum_{i=1}^{m}[S_i-\alpha]^+$ and $\frac{1}{m}\sum_{i=1}^{m}\rho_\epsilon(S_i-\alpha)$ becomes smaller. In addition the function $\frac{1}{m}\sum_{i=1}^{m}[S_i-\alpha]^+$ appears smoother.

Using $\tilde{F}(x,\alpha)$ as a continuously differentiable approximation to $F(x,\alpha)$, we solve the following continuous piecewise quadratic convex programming problem

$$\begin{array}{rl} & \min\limits_{(x,\alpha)} \left(\tilde{F}_\beta(x,\alpha)+\sum_{j=1}^{n}c_j|x_j|\right) \\ \text{subject to} & \left(V^0\right)^T x = 1 \\ & \left(\overline{\delta V}\right)^T x = \bar{r} \\ & l \le x \le u \end{array} \tag{29}$$

Note that, for (26), each simulation introduces an additional variable (and constraint) in its equivalent linear program formulation when the piecewise linear function $\bar{F}(x,\alpha)$ is used to approximate $F(x,\alpha)$. On the other hand, when the continuously differentiable function $\tilde{F}(x,\alpha)$ is used to approximate $F(x,\alpha)$, the minimization problem (29) has $(n+1)$ independent variables and its equivalent nonlinear program formulation only has $O(n)$ independent variables and constraints.

An optimization method for a convex nonlinear programming problem (29) typically generates an infinite sequence of approximations converging to a solution. At each iteration, however, it typically requires a function and a gradient evaluation and $O(n^3)$ linear algebraic operations. The function/gradient evaluation costs $O(mn)$. If exact second order derivatives

	MOSEK (cpu sec)			Smoothing (cpu sec)		
	# of instruments being considered					
# scenarios	8	48	200	8	48	200
10000	6.47	42.04	4244.30	2.45	16.78	419.52
25000	33.50	98.91	10784.10	5.37	35.48	838.15
50000	36.01	318.72	-	9.90	62.08	2080.16

Table 5: CPU times for MOSEK vs. Smoothing: $\beta = 0.99$

are used, then the Hessian calculation in the worst case is $O(n^2 k)$ where k is the total number of simulations with $|-(\delta V)_i^T x - \alpha| \leq \epsilon$. Given that CVaR optimization minimizes the tail loss with a typical confidence level of $\beta \geq 0.9$, k is usually very small relative to m for most iterations.

Table 5 makes a comparison between the cpu times of the proposed smoothing formulation and the linear programming approach (interior point method software MOSEK is used here) for individual problems. We consider the derivative portfolio CVaR optimization problem whose investment universe consists of vanilla call and put options on the same four correlated assets described in §4 with the strikes and maturities described in Table 12 in Appendix B. The implementation of the smoothing method is based on an interior point method [5] for nonlinear minimization with bound constraints and is implemented in MATLAB v6.1. The comparison is made on a Sun Sparc Ultra-5_10 machine. We observe that the smoothing method is much more efficient than the linear programming approach with up to a 1187% efficiency speedup. In addition, the 200 instruments and 50000 simulations example can now be solved in less than 35 cpu minutes with the smooth formulation due to less memory requirement and better computational efficiency.

Next we illustrate the accuracy and computational efficiency of this smoothing technique in greater detail. We now consider a different set of derivative portfolios on the same four correlated assets described in §4. The portfolios consist of an equal number of vanilla calls, vanilla puts, binary calls and binary puts on each asset. Once again the options are specified by all combinations of strikes $(K_n \times S_0)$ and expiry times $(T_n \times \bar{t})$, where n is the number of instruments (description in Table 12 in Appendix B.) The investment horizon we use here is $\bar{t} = 62.5$ days.

For various costs, Table 6 compares the CPU usage of MOSEK and our smoothing technique. We observe that the smoothing technique is more efficient compared to the linear programming method software MOSEK; the best cpu efficiency speedup is achieved with no cost consideration. However, with a larger parameter of $\omega = 0.01$, MOSEK requires less cpu than when $\omega = 0.005$, possibly due to the improved conditioning of the problem for a larger ω.

For comparison, we consider the relative difference in risks Q_{VaR} and Q_{CVaR}, where Q_{VaR} is defined as

$$Q_{\text{VaR}} \stackrel{\text{def}}{=} \frac{\text{VaR}_s - \text{VaR}_m}{|\text{VaR}_m|}, \tag{30}$$

where VaR_s and VaR_m are the VaR values computed by the smoothing technique and MOSEK respectively. Note that we report CVaR as well as VaR since minimizing CVaR

m	n	$\omega = 0$		$\omega = 0.005$		$\omega = 0.01$	
		MOSEK	SMTH	MOSEK	SMTH	MOSEK	SMTH
	20	49.60	10.57	49.97	14.91	48.31	14.62
25000	100	826.95	92.58	885.72	267.14	687.96	177.41
	196	7484.89	875.29	4141.79	851.79	2258.44	1088.83
	20	129.67	24.40	120.99	35.15	124.37	47.16
50000	100	2893.16	182.08	1242.03	449.98	1068.60	412.04
	196	-	1413.31	-	1672.05	-	1545.49

Table 6: CPU times for varying costs and problem sizes, $\beta = 0.95$, $\bar{t} = 62.5$ days, and resolution parameter $\epsilon = 0.005$

m	n	$\omega = 0$		$\omega = 0.005$	
		$Q_{\text{VaR}}(\%)$	$Q_{\text{CVaR}}(\%)$	$Q_{\text{VaR}}(\%)$	$Q_{\text{CVaR}}(\%)$
	20	-0.6701	0.2114	-0.7111	0.1984
25000	100	-0.7007	0.9433	-1.3946	0.2442
	196	-1.1442	1.4990	-0.6874	0.1637
	20	-0.5598	-0.2362	-0.4370	0.0520
50000	100	-0.7407	0.3438	-0.8173	0.1572

Table 7: Comparison of VaR/CVaR values computed by MOSEK and the proposed smoothing technique for $\beta = 0.95$, $\bar{t} = 62.5$ days, and resolution parameter $\epsilon = 0.005$

typically leads to a small VaR and VaR itself is an important risk measure. It is implicitly assumed that the number of scenarios and ω are fixed in computing Q_{VaR}. The quotient Q_{CVaR} is defined in a similar manner.

Table 7 compare risks of the optimal portfolios computed by linear programming method software MOSEK and the smoothing method for different weighted cost parameters; we observe that the relative difference is less than 1.5% for this example.

For smoothing method, the continuously differentiable approximation $\tilde{F}(x, \alpha)$ to $F(x, \alpha)$ depends on the resolution parameter ϵ. Typically this parameter is set to have a value between 0.05 to 0.005; the resolution parameter value should be smaller for a larger number of simulations since this leads to better approximation. The resolution parameter of 0.005 typically leads to a negligible difference in optimal risks between the portfolios computed from the linear programming method and the smoothing method. Table 8 illustrates the effect of the resolution parameter ϵ, on the cpu requirement and relative risk difference to that computed from MOSEK. The resolution parameters used here are $\epsilon = 0.005,\ 0.001,$ and 0.0005. We make a few interesting observations. Firstly, the risks computed from the smoothing method can be smaller than those computed by MOSEK; this suggests that the smooth approximation is an acceptable approximation to the augmented CVaR function, if not more preferable. Secondly, as ϵ becomes smaller, the relative risk difference becomes smaller.

$\epsilon = 0.005$				
m	n	$Q_{\text{VaR}}(\%)$	$Q_{\text{CVaR}}(\%)$	CPU Time
25000	20	-0.7348	0.1856	14.62
	100	-0.5261	0.8342	177.41
	196	-0.5941	1.7078	1088.83
50000	20	-0.4684	0.0385	47.16
	100	-1.0684	-0.1008	412.04
	196	-	-	1545.49

$\epsilon = 0.001$				
m	n	$Q_{\text{VaR}}(\%)$	$Q_{\text{CVaR}}(\%)$	CPU Time
25000	20	-0.1307	-0.0008	21.06
	100	-0.2883	0.0398	203.87
	196	-0.2057	0.0269	553.14
50000	20	-0.0521	-0.0016	44.43
	100	-0.0155	-0.0445	487.08
	196	-	-	1621.86

$\epsilon = 0.0005$				
m	n	$Q_{\text{VaR}}(\%)$	$Q_{\text{CVaR}}(\%)$	CPU Time
25000	20	-0.0005	0.0000	33.61
	100	-0.0015	0.0001	238.13
	196	-0.0028	0.0007	647.71
50000	20	-0.0009	-0.0002	48.34
	100	-0.0051	-0.0012	449.56
	196	-	-	1720.36

Table 8: Comparison of VaR/CVaR values computed by MOSEK and the proposed smoothing technique for different resolution parameters ϵ, $\omega = 0.01$, $\beta = 0.95$, and $\bar{t} = 62.5$ days

6 Concluding Remarks

In this paper we analyze the well-posedness of the derivative portfolio risk minimization problem with CVaR and VaR as the choice of risk measures. We illustrate that this minimization problem is typically ill-posed for derivative portfolios. In particular, we have shown that, when the derivative values are computed through delta-gamma approximations, there typically are an infinite number of portfolios with the same VaR and CVaR. Thus the derivative portfolio selection problem of minimizing risk subject to a specified return typically has an infinite number of solutions when using delta-gamma approximations. When the derivative values are computed using more accurate methods such as Black-Scholes formulae and Monte Carlo techniques, the optimal CVaR or VaR problem is typically ill-posed.

We illustrate that one may not be able to remove the ill-posedness of the CVaR/VaR optimization problem by simply adding constraints. When simple bound constraints are imposed on the instrument holdings, the optimal CVaR portfolio typically has a large number of non-zero instrument holdings (mostly at their bounds). This type of optimal portfolio may not be desirable and can be problematic since it may entail large management and transaction costs, and it tends to magnify modeling error.

We propose the inclusion of a weighted cost consideration in the
CVaR optimization problem. We model the cost as proportional to the magnitude of instrument holding; this cost model is capable of controlling transaction cost as well as management cost. We illustrate that minimizing CVaR together with this cost model leads to more desirable portfolios with significantly smaller transaction costs, fewer non-zero instrument holdings, and comparable CVaR (and VaR) measures.

We propose a computationally efficient method for solving a simulation based CVaR optimization problem by exploiting the fact that the objective function in the CVaR optimization problem approaches a continuously differentiable function as the number of Monte Carlo samples increases to infinity. With a preliminary implementation of the proposed method in MATLAB, a comparison is made with the standard linear programming approach. We illustrate that solving a continuously differentiable piecewise quadratic approximation to the CVaR optimization problem is much more efficient, producing an optimal CVaR, for appropriately chosen resolution parameters, very close to that obtained without cost consideration. Furthermore, it is more suitable for solving large scale CVaR portfolio optimization problems.

Although we have focused, in this paper, on the optimal derivative portfolio investment problem for both theoretical analysis and computational illustrations, similar analysis and computational results on the effectiveness of the CVaR and cost minimization formulation are presented in [2] for derivative portfolio hedging problems.

References

[1] *The MOSEK Optimization Tools Version 2.5: User's manual and reference*. Mosek Aps, Denmark, 2002.

[2] S. Alexander, T. F. Coleman, and Y. Li. Derivative portfolio hedging based on CVaR. In G. Szegö, editor, *Risk Measures for the 21st Century*, 339–363, Wiley, 2003.

[3] P. Artzner, F. Eber, J.M. Eber, and D. Heath. Thinking coherently. *Risk*, 10:68–71, 1997.

[4] P. Artzner, F. Eber, J.M. Eber, and D. Heath. Coherent measures of risk. *Mathematical Finance*, 9:203–228, 1999.

[5] T. F. Coleman and Y. Li. An interior, trust region approach for nonlinear minimization subject to bounds. *SIAM Journal on Optimization*, 6(2):418–445, 1996.

[6] J. Cvitanic and I. Karatzas. Hedging and portfolio optimization under transaction costs: a martingale approach. *Mathematical Finance* 6:133–165, 1996.

[7] G.B. Dantzig. *Linear Programming and extensions*. Princeton University Press, 1963.

[8] M.H.A. Davis and A.R. Norman. Portfolio selection with transaction costs,. *Mathematics for Operations Research* 15:676–713, 1990.

[9] D. Duffie and J. Pan. An overview of value at risk. *Journal of Derivatives*, pages 7–49, Spring 1997.

[10] R. Fletcher. *Practical Methods of Optimization: Volume 2, Constrained Optimization*. John Wiley and Sons, 1981.

[11] J. Hadamard. *Lectures on Cauchy's Problem in Linear Partial Differential Equations*. Yale University Press, New Haven, 1923.

[12] N. Karmarkar. A new polynomial-time algorithm for linear programming. *Combinatorica*, 4:373–395, 1984.

[13] H. Konno and A. Wijayanayake. Mean-Absolute Deviation Portfolio Optimization Model under Transaction Costs. Department of IE and Management, Tokyo Institute of Technology, 1999.

[14] R. Mansini, W. Ogryczak and M.G. Speranza. LP Solvable Models for Portfolio Optimization: A Classification and Computational Comparison. *IMA Journal of Management Mathematics*, 2003, to appear.

[15] H. Mausser and D. Rosen. Beyond VaR: From measuring risk to managing risk. *ALGO Research Quaterly*, 1:5–20, 1999.

[16] R. McKay and T.E. Keefer. VaR is a dangerous technique. *Corporate Finance Searching for Systems Integration Supplement*, September:30, 1996.

[17] W. Ogryczak and A. Ruszczynski. Dual Stochastic Dominance and Quantile Risk Measures. *International Transactions in Operational Research*, 9:661–680, 2002.

[18] G.Ch Pflug. Some remarks on the value-at-risk and the conditional value-at-risk. In S. Uryasev, editor, *Probablistic Constrained Optimization: Methodology and Applications*. Kluwer Academic Publishers, 2000.

[19] R. T. Rockafellar and S. Uryasev. Optimization of conditional value-at-risk. *Journal of Risk*, 2(3):21–41, 2000.

[20] R. T. Rockafellar and S. Uryasev. Conditional value-at-risk for general loss distributions. *Journal of Banking and Finance*, 26(7):1443–1471, 2002.

Appendix

A CVaR and Delta-Gamma Approximation

In §3 we assumed that each instrument depended on a single risk factor. In general, an instrument value may depend on more than one risk factor. Assuming there are d risk factors, in the analysis that follows we discuss the conditioning of minimizing derivative portfolio VaR/CVaR for this general case. We first need to set up a matrix of second order sensitivities of derivatives to the underlying risk factors. Assume that for each h, $1 \leq h \leq n$, V_h depends on r_h risk factors, where $r_h \leq d$. Then the second order sensitivity matrix Γ_h, corresponding to V_h, has at most r_h^2 non-zero entries. Further, since Γ_h is symmetric, it has at most $\frac{r_h(r_h+1)}{2}$ distinct non-zero entries. Assuming an ordering on the risk factors, construct the sets $R_h = \{(i,j)|V_h \text{ is dependent on risk factors } i \text{ and } j, i.e., \Gamma_h(i,j) \neq 0, i \leq j\}$ for $h = 1, \ldots, n$. Then

$$(\delta S)^T \Gamma_h (\delta S) = \sum_{(i,i)\in R_h} \Gamma_h(i,i)\delta S_i^2 + \sum_{(i,j)\in R_h, i\neq j} 2\Gamma_h(i,j)\delta S_i \delta S_j$$

Then consider $\hat{R} = \bigcup_{h=1}^{n} R_h$. We also set an order on $\hat{R}$ as follows: for $(i_1, j_1), (i_2, j_2) \in \hat{R}$,

$$\begin{array}{ll} (i_1, j_1) < (i_2, j_2) & \text{if } i_1 < i_2, \\ (i_1, j_1) < (i_2, j_2) & \text{if } i_1 = i_2 \text{ and } j_1 < j_2. \end{array}$$

Now we construct the second order sensitivity matrix $\Gamma \in \mathbb{R}^{\hat{d}\times n}$ for all instruments under consideration, where $\hat{d}$ is the cardinality of $\hat{R}$. Row i of Γ corresponds to the i^{th} smallest element of $\hat{R}$ which we shall refer to as $\hat{R}_i$. Now

$$\Gamma(l,h) \stackrel{\text{def}}{=} \begin{cases} 0 & \text{if } \hat{R}_l \notin R_h \\ \frac{\partial^2 V_h^0}{\partial S_i^2}, \text{ where } \hat{R}_l = (i,i) & \text{if } \hat{R}_l \in R_h \\ 2\frac{\partial^2 V_h^0}{\partial S_i \partial S_j}, \text{ where } \hat{R}_l = (i,j) & \text{if } \hat{R}_l \in R_h, i \neq j \end{cases} \tag{31}$$

We also construct the vector $\delta\hat{S}^2 \in \mathbb{R}^{\hat{d}}$ which represents the required corresponding second order change in risk factors, $(\delta\hat{S}^2)_l \stackrel{\text{def}}{=} \delta S_i \delta S_j$, where $\hat{R}_l = (i,j)$.

THEOREM A.1. *Assume that a portfolio is formed from instruments $\{V_1, \cdots, V_n\}$ and the underlying risk factors of $\{V_1, \cdots, V_n\}$ are $\{S_1, \cdots, S_d\}$. For a fixed investment horizon $\bar{t} > 0$, $1 \leq i \leq n$, assume that*

$$V_i^{\bar{t}} - V_i^0 = \left(\frac{\partial V_i^0}{\partial t}\right)\delta\bar{t} + \left(\frac{\partial V_i^0}{\partial S}\right)^T (\delta S) + \frac{1}{2}(\delta S)^T \Gamma_i (\delta S).$$

Then the following is true:

1. *If $n > \hat{d} + d + 1$, where $\hat{d}$ is defined as above, and there exists a portfolio whose VaR = VaR* and CVaR =CVaR*, where VaR* and CVaR* are the minimal VaR and CVaR, then there are an infinite number of optimal portfolios that have the minimal VaR and CVaR.*

2. *If* $X = \{x : (V^0)^T x = 1,\ \left(\overline{\delta V}\right)^T x = \bar{r}\}$ *and* $n > \hat{d} + d + 3$*, then the optimal CVaR and VaR portfolios defined by* $\min_{x \in X} \phi_\beta(x)$ *and* $\min_{x \in X} \alpha_\beta(x)$*,*

   for any $0 < \beta < 1$*, lie in a linear subspace of dimension* $n - (\hat{d} + d + 3)$*.*

Proof. The results in the theorem hold due to the fact that there are an infinite number of portfolios with the same VaR and CVaR under the assumed assumption.

We first note that if portfolios $x^{(1)}$, $x^{(2)} \in \mathbb{R}^n$ satisfy

$$f(x^{(1)}, S) \equiv f(x^{(2)}, S), \quad \text{for all possible } S,$$

then these two portfolios have the same VaR value and the same CVaR value.

Let

$$\Lambda \stackrel{\text{def}}{=} \left[\left(\frac{\partial V^0}{\partial t} \right),\ \left(\frac{\partial V^0}{\partial S} \right),\ \frac{1}{2}\Gamma \right] \in \mathbb{R}^{n \times (\hat{d}+d+1)}. \tag{32}$$

The proof is straight forward from the observation that

$$\begin{aligned} f(x, S) &= -x^T \left(\left(\frac{\partial V^0}{\partial t} \right) \delta \bar{t} + \left(\frac{\partial V^0}{\partial S} \right) (\delta S) + \frac{1}{2}\Gamma \left(\delta \hat{S}^2 \right) \right) \\ &= -x^T \Lambda \begin{bmatrix} \delta \bar{t} \\ \delta S \\ (\delta \hat{S}^2) \end{bmatrix} \end{aligned}$$

If $n > \hat{d} + d + 1$, there exists a non-zero $z \in \mathbb{R}^n$ satisfying

$$\Lambda^T z = 0.$$

Then $f(x + \theta z, S) = f(x, S)$ for any S and θ.

For the second result, similarly there exists a non-zero $z \in \mathbb{R}^n$ which lies in the null space of $\left[\Lambda,\ V^0,\ \overline{\delta V}\right]^T$ (this null space has dimension $n - (\hat{d} + d + 3)$).

Note that in the worst case, $\hat{d} = \frac{d(d+1)}{2}$. This completes our proof. □

B Data Specifications for Computational Examples

Tables 9, 10, and 11 describe the initial underlying prices, the expected annual rates of return, and the covariance matrix of the rate of return respectively.

100	50	30	100

Table 9: Initial Asset Prices

0.1091	0.0619	0.0279	0.0649

Table 10: Annual Expected Rate of Return

0.2890	0.0690	0.0080	0.0690
0.0690	0.1160	0.0200	0.0610
0.0080	0.0200	0.0220	0.0130
0.0690	0.0610	0.0130	0.0790

Table 11: Annual Covariance Matrix

Tables 12-14 describe the strike prices and the expiry of various portfolios used in the paper for different investment horizons; they are classified according to the total number of instruments in the portfolio universe.

8 instrument example	
K_8	1
T_8	4

48 instrument example			
K_{48}	0.8	1	1.25
T_{48}	2	4	

200 instrument example					
K_{200}	0.8	0.9125	1.025	1.1375	1.25
T_{200}	2	3.5	5	6.5	8

20 instrument example (underlying assets included)	
K_{20}	1
T_{20}	4

100 instrument example (underlying assets included)			
K_{100}	0.9	1	1.1
T_{100}	3	6	

196 instrument example (underlying assets included)				
K_{196}	0.8	1.025	1.25	
T_{196}	2	4	6	8

Table 12: Strike Price Equals $K_n \times S_0$ and Expiries Equal $T_n \times \bar{t}$ for $\bar{t} = 10$ days

20 instrument example (underlying assets included)	
K_{20}	1
T_{20}	2

100 instrument example (underlying assets included)			
K_{100}	0.9	1	1.1
T_{100}	2	4	

196 instrument example (underlying assets included)				
K_{196}	0.8	1.025	1.25	
T_{196}	1.5	2	3	4

Table 13: Strike Price Equals $K_n \times S_0$ and Expiries Equal $T_n \times \bar{t}$ for $\bar{t} = 62.5$ days

20 instrument example (underlying assets included)		
K_{20}	1	
T_{20}	2	

100 instrument example (underlying assets included)			
K_{100}	0.9	1	1.1
T_{100}	2	4	

196 instrument example (underlying assets included)				
K_{196}	0.8	1.025	1.25	
T_{196}	1.5	2	3	4

Table 14: Strike Price Equals $K_n \times S_0$ and Expiries Equal $T_n \times \bar{t}$ for $\bar{t} = 125$ days