



ELSEVIER

Available online at www.sciencedirect.com

SCIENCE @ DIRECT®

Journal of Banking & Finance xxx (2005) xxx–xxx

**Journal of
BANKING &
FINANCE**

www.elsevier.com/locate/jbf

Efficient fund of hedge funds construction under downside risk measures

David P. Morton ^a, Elmira Popova ^b, Ivilina Popova ^{c,*}

^a Graduate Program in Operations Research and Industrial Engineering, The University of Texas at Austin,
Austin, TX 78712, USA

^b Graduate Program in Operations Research and Industrial Engineering, The University of Texas at Austin,
Austin, TX 78712, USA

^c Albers School of Business and Economics, Seattle University, Seattle, WA 98122, USA

Abstract

We consider portfolio allocation in which the underlying investment instruments are hedge funds. We consider a family of utility functions involving the probability of outperforming a benchmark and expected regret relative to another benchmark. Non-normal return vectors with prescribed marginal distributions and correlation structure are modeled and simulated using the normal-to-anything method. A Monte Carlo procedure is used to obtain, and establish the quality of, a solution to the associated portfolio optimization model. Computational results are presented on a problem in which we construct a fund of 13 CSFB/Tremont hedge-fund indices.

© 2005 Elsevier B.V. All rights reserved.

JEL classification: G11; C15; C61

Keywords: Portfolio choice; Expected regret; Hedge funds; Fund of funds; Portfolio optimization; Monte Carlo simulation

* Corresponding author.

E-mail addresses: morton@mail.utexas.edu (D.P. Morton), elmira@mail.utexas.edu (E. Popova), ipopova@austin.rr.com (I. Popova).

1. Introduction

Over the last decade, investors have faced a new investment opportunity that is difficult to map to any existing asset class. So-called alternative investments are usually available only to high net-worth individuals or institutional investors; they vary widely in terms of styles, underlying strategies, and as a consequence, in performance. Hedge-fund returns usually differ substantially from returns of standard asset classes and there is growing evidence that, as a result, they may be of interest to investors (see, e.g., [Fung and Hsieh, 1997](#); [Siegmann and Lucas, 2002](#)). Important issues include identifying the right proportion to invest in hedge funds and how to construct a portfolio of hedge funds. Studies of the statistical properties of hedge-fund indices argue that standard methods used for portfolio construction (e.g., mean–variance analysis) can be inadequate for hedge funds, e.g., [Agarwal and Naik \(2001\)](#), [Brooks and Kat \(2002\)](#), [Fung and Hsieh \(1997\)](#), [Fung and Hsieh \(2001\)](#), [Martin \(2001\)](#) and [Popova et al. \(2003\)](#). Usually, negative or positive skewness and larger kurtosis, relative to standard asset classes, are observed in monthly return series of hedge funds. Additionally, these return observations are typically not time-independent; instead positive autocorrelation is often observed. This leads to underestimation of the volatility, and as a result, overestimation of the Sharpe ratio, if it is used as a performance measure.

In this paper, we consider an investment problem in which the underlying assets are hedge funds. We focus on a family of utility functions, which incorporate benchmark returns. The performance of a portfolio manager is often judged relative to that of a (possibly random) benchmark. For example, a benchmark return for a mutual fund manager might be the S&P 500 Index or the Russell 3000 Index while benchmarks for bond managers include indices of Lehman Brothers and Salomon Brothers, among others. Perhaps the simplest performance measure relative to a benchmark is the binary outcome of whether the benchmark is achieved. For continuous-time trading strategies involving the probability of achieving a stochastic benchmark see [Browne \(1999\)](#). Another performance measure is the magnitude by which the benchmark is missed, on the downside. This leads to minimizing so-called expected regret ([Dembo and King, 1992](#)), relative to the benchmark. In Section 2, we discuss our family of expected utilities, which are a weighted sum of the probability of achieving a benchmark and expected regret relative to another benchmark. The former benchmark (e.g., the S&P 500 return) is chosen to be more aggressive than the latter (e.g., the risk-free rate). Due to the manner in which today's portfolio managers are judged, we regard the probability of achieving the benchmark as the reward measure and we use expected regret as the measure of risk.

In Section 3 we describe our use of the normal-to-anything (NORTA) method ([Cario and Nelson, 1997](#)) for modeling, and simulating from, non-normal return distributions. This approach takes as input marginal distributions for each asset's return, which need not be from the same parametric family, and the correlation between the returns for each pair of assets. It then transforms a multi-variate normal vector into a non-normal random vector, with the aim of preserving the prescribed marginal distributions and correlation matrix.

Under general return distributions, it is not possible to solve the associated portfolio optimization problems exactly. So, in Section 4 we describe a Monte Carlo approximation procedure for obtaining a solution. This method is quite natural and straightforward to apply, provided one can simulate, say, n observations from the return distribution and solve the associated portfolio optimization model with these n realizations of the return vector. King and Jensen (1992), for example, propose such a procedure to obtain an approximately-optimal solution to a portfolio optimization problem. A distinguishing feature of our approach is that it includes a method for establishing the quality of the recommended allocation decision, in terms of a confidence interval. In this paper, we use the NORTA method to simulate return observations and we use the benchmark-based utility function described above. That said, one of the attractive features of our solution approach is that it is not restricted to these choices, and it is our view that the proposed methodology may help investors address some of the issues that arise in optimizing hedge-fund allocation. In Section 5, our proposed methodology is illustrated by constructing a fund of 13 hedge-fund indices with the goal of outperforming a benchmark return of 10% per year.

2. A benchmark-based utility

Consider an investment problem with m underlying instruments having random returns $\tilde{\xi} = (\tilde{\xi}_1, \dots, \tilde{\xi}_m)$. We will construct a portfolio by choosing weights $x = (x_1, \dots, x_m)$ satisfying $x \in X = \{x : \sum_{i=1}^m x_i = 1, l \leq x \leq u\}$. Nominal simple bounds of $l_i = 0$ and $u_i = 1, i = 1, \dots, m$, may be tightened depending on the setting. The random return of our portfolio is $\tilde{R}_x = \tilde{\xi}x = \sum_{i=1}^m \tilde{\xi}_i x_i$, and we seek an allocation decision, x^* , which solves

$$z^* = \max_{x \in X} \left[Eu_\lambda(\tilde{R}_x) = P(\tilde{R}_x \geq \tilde{r}_1) - \lambda E(\tilde{r}_2 - \tilde{R}_x)^+ \right], \tag{1}$$

where

$$u_\lambda(\tilde{R}_x) = I(\tilde{R}_x \geq \tilde{r}_1) - \lambda(\tilde{r}_2 - \tilde{R}_x)^+. \tag{2}$$

The first term on the right-hand side of (2) involves the event of outperforming the possibly stochastic benchmark \tilde{r}_1 , and the second term is the down-side deviation from another possibly stochastic benchmark \tilde{r}_2 . The indicator function $I(\cdot)$ takes value one if its argument is true and zero otherwise, and $(\cdot)^+ = \max\{\cdot, 0\}$. We use $\tilde{\xi}$ and \tilde{R}_x to denote the random return vector and random portfolio value and ξ and R_x to denote their realizations.

Utility u_λ rewards outperformance of benchmark \tilde{r}_1 and with weighting factor $\lambda \geq 0$, penalizes down-side deviations from \tilde{r}_2 . We typically choose \tilde{r}_1 as the more aggressive benchmark, e.g., the return of an index fund such as the S&P 500 that we would like to outperform. Benchmark \tilde{r}_2 is more conservative, e.g., $\tilde{r}_2 = 0$ or the risk-free rate.

Model (1) has the following special cases. With $\lambda = 0$, it is $\max_{x \in X} P(\tilde{R}_x \geq \tilde{r}_1)$, i.e., we should select an allocation x^* that maximizes the probability of outperforming

benchmark \tilde{r}_1 . As λ grows, increasing weight is put on the expected regret objective and the limiting model is $\min_{x \in X} E(\tilde{r}_2 - \tilde{R}_x)^+$, i.e., to minimize expected regret, relative to benchmark \tilde{r}_2 . An equivalence between minimizing CVaR (Rockafellar and Uryasev, 2000; Rockafellar and Uryasev, 2002) and minimizing expected regret, for appropriate quantile-level and fixed benchmark pairs, is established in Testuri and Uryasev (2004).

We emphasize that we use the term “utility” for u_λ liberally. The first term in (2) is only concerned with the binary outcome of whether the benchmark is achieved and ignores the magnitude by which it is achieved or missed. As a result, u_λ is discontinuous and neither u_λ nor its expectation are concave functions. So, our utility function does not satisfy the notion of coherence (Artzner et al., 1997, 1999). However, to the extent that today’s portfolio managers are judged primarily by whether they outperform a benchmark, this utility may provide a reasonable model of their goals. Moreover, our paper has its roots in extensive discussions with institutional investors, mutual-fund managers, hedge-fund managers, and pension-plan sponsors, in which they viewed such performance measures as playing a central role. So, we regard the methods we develop here to solve this model as important and relevant to modern practice.

Following the spirit of Markowitz (1952), one commonly-used approach is to maximize a reward measure (e.g., ER_x) subject to $x \in X$ and the additional constraint that a risk measure (variance, CVaR, expected regret, etc.) is at most a prespecified level ℓ . Doing so for a range of values of ℓ is equivalent to optimizing a weighted sum of the reward and risk measures for a range of values of that weight. We are interested in benchmark-based utility functions for reasons outlined above. In utility u_λ we use expected regret as the measure of risk, and we replace the typical reward measure of mean return with the probability of achieving benchmark \tilde{r}_1 because it is our view that this more accurately captures how portfolio managers are motivated (rewarded). In our subsequent analysis we point to implications that can arise in this setting due to our non-standard utility.

3. Probability model and simulation

Given a data set of what is presumed to be independent and identically distributed (iid) observations of a univariate random variable, well-known methods for selecting a probability distribution, fitting its distributional parameters, and assessing the resulting goodness-of-fit are routinely practiced in a wide variety of application areas. In the context of random vectors, this approach is much more difficult to carry out, in part because of the size of data sets needed to select and fit multi-variate distributions but also because of associated methodological and computational issues. These difficulties are compounded as the dimension of the vector grows or when the marginal distributions are not of the same parametric family.

A circumvention is to specify univariate marginal distributions for each component of the random vector and to specify the pair-wise correlation structure of the random vector’s components. Then, the goal is to find a method to select a multi-

variate distribution that has the specified marginals and has, or nearly has, the specified correlation matrix.

This type of approach has shortcomings (but viable alternatives are limited): (i) In general, multiple multi-variate distributions can have the same marginal distributions and correlation matrix; (ii) it is possible that there is no joint distribution that has both the specified marginal distributions and the specified correlation matrix; (iii) a particular method may not be capable of generating a distribution that has the desired marginals and correlation structure, even though they are feasible in tandem.

Broadly speaking, (i) is recognized but is accepted as we only attempt to capture a first-order notion of dependence. Difficulty (ii) is addressed in Ghosh and Henderson (2002), where a computational method is given for testing whether specified marginals and correlation matrix are a feasible combination. Issue (iii) is discussed further below.

The approach we use here is the NORTA (normal-to-anything) method of Cario and Nelson (1997). It is designed to generate iid observations of a random vector $\tilde{\xi} = (\tilde{\xi}_1, \dots, \tilde{\xi}_m)$ with prescribed (product moment) correlation matrix $\Sigma_{\tilde{\xi}}$ and prescribed marginal distributions $F_i, i = 1, \dots, m$. Of course, this fits well in our setting because in the next section we employ a Monte Carlo approximation procedure for approximately solving model (1), which requires iid observations of $\tilde{\xi}$. In its original form, NORTA can suffer from (iii). That said, if one is willing to modify, slightly, the correlation matrix $\Sigma_{\tilde{\xi}}$ used as input then the NORTA method works, and an approach for doing so, involving semidefinite programming, is given in Ghosh and Henderson (2002, 2003), although heuristics may also be used.

For notational simplicity we are describing the NORTA method in terms of the return vector $\tilde{\xi}$, but if benchmarks \tilde{r}_1 and \tilde{r}_2 are random the method is actually performed, with obvious modifications, with respect to the random $(m + 2)$ -vector $(\tilde{\xi}, \tilde{r}_1, \tilde{r}_2)$. The NORTA procedure is summarized below.

- *Input:* Marginal distributions $F_i, i = 1, \dots, m$, and $m \times m$ correlation matrices $\Sigma_{\tilde{\xi}}$ and Σ_Z .
- *Output:* An observation of a random vector $\tilde{\xi}$ with marginals $F_i, i = 1, \dots, m$, and correlation matrix $\Sigma_{\tilde{\xi}}$.

1. Generate a multi-variate normal Z with mean zero, correlation matrix Σ_Z , and $\text{Var}Z_i = 1, i = 1, \dots, m$.
2. Form

$$\tilde{\xi} = \begin{pmatrix} F_1^{-1}[\Phi(Z_1)] \\ F_2^{-1}[\Phi(Z_2)] \\ \vdots \\ F_m^{-1}[\Phi(Z_m)] \end{pmatrix},$$

where Φ is the distribution function of a standard univariate normal, and $F_i^{-1}(x) = \inf\{u : F_i(u) \geq x\}, i = 1, \dots, m$.

By construction, $\Phi(Z_i)$ is a continuous uniform random variable on $(0, 1)$ and hence the distribution of $\tilde{\xi}_i = F_i^{-1}[\Phi(Z_i)]$ matches the specified marginal F_i . (This is the well-known “inversion” method of generating non-uniform random variates (Law and Kelton, 2000, Section 8.2.1).)

The correlation matrix Σ_Z is labeled as input to the above procedure. However, selecting Σ_Z is the main issue in carrying out the NORTA method, and it is done in a preprocessing step. The goal is to select Σ_Z so that the vector $\tilde{\xi}$ formed in step 2 has the desired correlation matrix $\Sigma_{\tilde{\xi}}$. To see how to do so, note that

$$\text{Corr}(\tilde{\xi}_i, \tilde{\xi}_j) = \frac{E[\tilde{\xi}_i \tilde{\xi}_j] - E[\tilde{\xi}_i]E[\tilde{\xi}_j]}{\sqrt{\text{Var}[\tilde{\xi}_i]\text{Var}[\tilde{\xi}_j]}}. \tag{3}$$

The first two moments of each $\tilde{\xi}_i$ are specified by its marginal distribution F_i , and so the only degree of freedom in $\text{Corr}(\tilde{\xi}_i, \tilde{\xi}_j)$ is in the cross-moment $E[\tilde{\xi}_i \tilde{\xi}_j]$, which is defined by

$$E[\tilde{\xi}_i \tilde{\xi}_j] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F_i^{-1}[\Phi(z_i)]F_j^{-1}[\Phi(z_j)]\phi_{ij}(z_i, z_j) dz_i dz_j, \tag{4}$$

where ϕ_{ij} is the bivariate normal density of (Z_i, Z_j) , i.e., with standard normal marginals and correlation coefficient $\rho_{ij} \equiv \text{Corr}(Z_i, Z_j)$. The correlation coefficient $\Sigma_{\tilde{\xi}}(i, j)$ is a numerical value given as input to the NORTA procedure. The correlation coefficient $\text{Corr}(\tilde{\xi}_i, \tilde{\xi}_j)$ is defined by the specified marginals F_i and F_j and the correlation coefficient ρ_{ij} . We can numerically solve the equation

$$\Sigma_{\tilde{\xi}}(i, j) = \text{Corr}(\tilde{\xi}_i, \tilde{\xi}_j) \tag{5}$$

by adjusting the value of ρ_{ij} , which determines $\text{Corr}(\tilde{\xi}_i, \tilde{\xi}_j)$ through (3) and (4). Doing so for the entire correlation matrix involves $m(m - 1)/2$ separate univariate root-finding problems, i.e., each ρ_{ij} , $1 \leq i < j \leq m$, is adjusted separately in order to enforce (5). The right-hand side of (5) is a non-decreasing function of ρ_{ij} (Cario and Nelson, 1997), which aids the root-finding process.

So, in order to run the NORTA procedure we first run a preprocessing step to compute Σ_Z . Let $\Lambda_Z(i, j) = \rho_{ij}$, $1 \leq i < j \leq m$, denote the coefficients found by the above numerical root-finding procedure. Assume, for the moment, that the matrix Λ_Z obtained in this way is positive semidefinite. Then, the correlation matrix $\Sigma_Z = \Lambda_Z$ is used as input to the NORTA procedure and the more detailed substeps needed to execute step 1 are:

- 1a. Perform a Cholesky factorization of Σ_Z so that we have C with $CC^T = \Sigma_Z$.
- 1b. Generate W , a random m -vector whose components are iid standard normal random variables.
- 1c. Form $Z = CW$.

The resulting vector Z is that used in step 1 of the NORTA procedure. Note that the Cholesky factorization of the covariance matrix is that which should be used in generating multi-variate normals, e.g., Fishman (1996, pp. 223–224). In our case, the

variances of the components of Z are unit-valued, and so the correlation matrix and covariance matrix of Z coincide.

Unfortunately, there is no guarantee that the matrix A_Z formed from the $m(m - 1)/2$ root-finding problems is positive semidefinite. If A_Z is positive semidefinite, we use $\Sigma_Z = A_Z$ as input to the NORTA method and proceed as described above. If A_Z is indefinite then it is obviously not a correlation matrix and cannot be used as input to NORTA. (This would be discovered if A_Z were blindly used in the above procedure when attempting the Cholesky factorization in step 1a.) Assuming that the specified correlation matrix Σ_ξ and marginals F_i , $i = 1, \dots, m$, are a feasible combination then if A_Z is indefinite NORTA is said to have failed, i.e., we are in case (iii) discussed at the beginning of this section. In this situation we follow the suggestion of Ghosh and Henderson (2003) by finding a positive semidefinite matrix Σ_Z , which minimizes $\|\Sigma_Z - A_Z\|$ (using e.g., the 1-norm). This requires solution of a semidefinite program, which we solve heuristically. The approach is justified by the fact that if Σ_Z is close to A_Z then the induced correlation matrix will be close to the desired Σ_ξ (Ghosh and Henderson, 2002).

In our computational work in Section 5 we use iid observations of $\tilde{\xi}$ generated using the NORTA method in which the correlation matrix Σ_ξ is the historical (empirical) correlation matrix based on $N = 108$ months of return data. For the marginal distributions F_i , $i = 1, \dots, m$, we use the *continuous* empirical distribution. In particular, given the sorted data for asset i , $\xi_i^{(1)} \leq \xi_i^{(2)} \leq \dots \leq \xi_i^{(N)}$, the continuous empirical distribution F_i is given by (e.g., (Law and Kelton, 2000, Section 6.2.4)):

$$F_i(x) = \begin{cases} 0 & \text{if } x < \xi_i^{(1)}, \\ \frac{j-1}{N-1} + \frac{x - \xi_i^{(j)}}{(N-1)(\xi_i^{(j+1)} - \xi_i^{(j)})} & \text{if } \xi_i^{(j)} \leq x < \xi_i^{(j+1)}, 1 \leq j \leq N-1, \\ 1 & \text{if } x \geq \xi_i^{(N)}. \end{cases}$$

4. Optimization and solution quality

When $\tilde{\xi}$ has a discrete distribution with a modest number of realizations we can solve the portfolio allocation problem (1) exactly (or, more precisely, within relatively tight numerical tolerances). In this case, let the realizations of $\tilde{\xi}$ be denoted ξ^ω , $\omega \in \Omega$, with probability mass function $p^\omega = P(\tilde{\xi} = \xi^\omega)$, $\omega \in \Omega$. Then (1) can be rewritten as the following mixed-integer linear program

$$\max_{x \in X, v, y} \sum_{\omega \in \Omega} p^\omega y^\omega - \lambda \sum_{\omega \in \Omega} p^\omega v^\omega, \tag{6a}$$

$$\text{s.t. } \zeta^\omega x \geq r_1^\omega y^\omega - M^\omega (1 - y^\omega), \omega \in \Omega, \tag{6b}$$

$$v^\omega \geq r_2^\omega - \zeta^\omega x, \omega \in \Omega, \tag{6c}$$

$$y^\omega \in \{0, 1\}, v^\omega \geq 0, \omega \in \Omega.$$

Binary decision variable y^ω indicates whether ($y^\omega = 1$) or not ($y^\omega = 0$) benchmark r_1^ω is achieved under scenario (sample point) $\omega \in \Omega$. To accomplish this the value of datum parameter M^ω is selected so that constraint (6b) is vacuous if $y^\omega = 0$. (The value $M^\omega = -\min_{i=1, \dots, m} \zeta_i^\omega$ suffices.) Continuous non-negative decision variables v^ω account for the magnitude of the shortfall of return relative to benchmark r_2^ω , in scenario ω , through constraint (6c). The objective function (6a) simply computes the weighted sum of the probability of shortfall and expected regret relative to these two benchmarks. We have indexed r_1^ω and r_2^ω by ω because they may be stochastic.

For general return distributions in which Ω has a large number of realizations, or is not finite, it can be difficult or impossible to solve (1) exactly. When the distribution of $\tilde{\zeta}$ is governed by the NORTA method with continuous empirical marginals, as described in Section 3, we are in the latter case of non-finite Ω because $\tilde{\zeta}$ is continuous. So, we resort to an approximation in which the expected utility, $Eu_\lambda(\tilde{\zeta}x)$, is replaced with a sample mean estimate. In particular, let $(\tilde{\zeta}^i, \tilde{r}_1^i, \tilde{r}_2^i)^1, \dots, (\tilde{\zeta}^i, \tilde{r}_1^i, \tilde{r}_2^i)^n$ be iid from the distribution of $(\tilde{\zeta}, \tilde{r}_1, \tilde{r}_2)$. We replace (1) with the following approximation

$$z_n^* = \max_{x \in X} \left[\frac{1}{n} \sum_{i=1}^n u_\lambda(\tilde{\zeta}^i x) = \frac{1}{n} \sum_{i=1}^n I(\tilde{\zeta}^i x \geq \tilde{r}_1^i) - \lambda \frac{1}{n} \sum_{i=1}^n (\tilde{r}_2^i - \tilde{\zeta}^i x)^+ \right]. \tag{7}$$

We regard this approximation as justified provided it possesses desirable large sample-size properties. Consistency results are available through the theory of epi- (or hypo-) convergence. The following is a special case of a theorem of Attouch and Wets (1990, Theorem 4.3) which directly applies to (1) and (7).

Theorem 1. *Let $\tilde{\eta}^1, \dots, \tilde{\eta}^n$ be iid as $\tilde{\eta}$ and let $u_0(x, \eta)$ be upper semicontinuous in x for all η . If there exists $\gamma \in \Re$ with $u_0(x, \tilde{\eta}) \leq \gamma$, with probability one (wp1) then $n^{-1} \sum_{i=1}^n u_0(x, \tilde{\eta}^i)$ hypo-converges to $Eu_0(x, \tilde{\eta})$, wp1.*

Our utility $u_0(x, (\zeta, r_1, r_2)) = I(\zeta x \geq r_1) - \lambda(r_2 - \zeta x)^+$ satisfies the conditions of Theorem 1 with $\gamma = 1$ since $\lambda \geq 0$. Let x_n^* be an optimal solution of (7), and recall that z^* and z_n^* are the optimal values of (1) and (7), respectively. The hypo-convergence property, together with the fact that the constraint set X is compact, implies consistency of z_n^* and x_n^* , i.e., z_n^* converges to z^* , wp1, and all limit points of $\{x_n^*\}_{n=1}^\infty$ solve (1), wp1. Existence of at least one limit point of $\{x_n^*\}_{n=1}^\infty$ in X is ensured as X is compact.

Computationally, we carry out the optimization in (7) by solving the mixed-integer program (6a) under the distribution with realizations $\tilde{\zeta}^1, \dots, \tilde{\zeta}^n$, sample space

$\Omega = \{1, \dots, n\}$ and probability mass function $p^\omega = 1/n$, $\omega \in \Omega$. While we are assured (via Theorem 1) that as n grows large the solutions of (7) approach optimal solutions, this tells us little about the quality of a feasible candidate allocation decision, x_n^* , obtained by solving an instance of (7) for finite n . In this regard, our goal is the following: Given $\hat{x} \in X$ (e.g., $\hat{x} = x_n^*$) we seek a confidence interval on the so-called optimality gap, $z^* - Eu_\lambda(\tilde{\xi}\hat{x})$, of the form

$$P\{z^* - Eu_\lambda(\tilde{\xi}\hat{x}) \leq \tilde{\epsilon}\} \approx 1 - \alpha, \tag{8}$$

where $\tilde{\epsilon}$ is a (random) confidence interval width. From (8) we will infer that while \hat{x} may be suboptimal, i.e., $z^* - Eu_\lambda(\tilde{\xi}\hat{x}) \geq 0$, the optimality gap at \hat{x} is at most $\tilde{\epsilon}$ with (approximate) probability $1 - \alpha = 0.95$ (say). So, if $\tilde{\epsilon}$ is sufficiently small we regard \hat{x} as being of high quality. We summarize our approach to finding a candidate portfolio allocation decision and establishing its quality in the following procedure, which is based on Mak et al. (1999) and Morton and Wood (1998).

- *Step 0:* Select sample sizes n_u, n_l and n and $0 < \alpha < 1$.
- *Step 1:* For $i = 1, \dots, n_u$:
 - (a) Generate $(\tilde{\xi}, \tilde{r}_1, \tilde{r}_2)^{i1}, \dots, (\tilde{\xi}, \tilde{r}_1, \tilde{r}_2)^{in}$ iid as $(\tilde{\xi}, \tilde{r}_1, \tilde{r}_2)$.
 - (b) Let (x_n^{*i}, z_n^{*i}) solve $\max_{x \in X} [\frac{1}{n} \sum_{j=1}^n I(\tilde{\xi}^{ij} x \geq \tilde{r}_1^{ij}) - \lambda \frac{1}{n} \sum_{j=1}^n (\tilde{r}_2^{ij} - \tilde{\xi}^{ij} x)^+]$.
- *Step 2:* Form $\bar{U}(n_u) = \frac{1}{n_u} \sum_{i=1}^{n_u} z_n^{*i}$ and $s_u^2(n_u) = \frac{1}{n_u-1} \sum_{i=1}^{n_u} (z_n^{*i} - \bar{U}(n_u))^2$.
- *Step 3:* Form candidate solution $\hat{x} = \frac{1}{n_u} \sum_{i=1}^{n_u} x_n^{*i}$.
- *Step 4:* Generate n_l iid observations, $(\tilde{\xi}, \tilde{r}_1, \tilde{r}_2)^{1l}, \dots, (\tilde{\xi}, \tilde{r}_1, \tilde{r}_2)^{n_l l}$. Let $\bar{L}(n_l) = \frac{1}{n_l} \sum_{i=1}^{n_l} I(\tilde{\xi}^i \hat{x} \geq \tilde{r}_1^i) - \lambda \frac{1}{n_l} \sum_{i=1}^{n_l} (\tilde{r}_2^i - \tilde{\xi}^i \hat{x})^+$ and $s_l^2(n_l) = \frac{1}{n_l-1} \sum_{i=1}^{n_l} (I(\tilde{\xi}^i \hat{x} \geq \tilde{r}_1^i) - \lambda(\tilde{r}_2^i - \tilde{\xi}^i \hat{x})^+ - \bar{L}(n_l))^2$.
- *Step 5:* To evaluate the quality of \hat{x} , form the approximate $(1 - \alpha)$ -level confidence interval $[0, [\bar{U}(n_u) - \bar{L}(n_l)]^+ + \tilde{\epsilon}_u + \tilde{\epsilon}_l]$, where $\tilde{\epsilon}_u = s_u(n_u)t_{n_u-1, \alpha/2} / \sqrt{n_u}$ and $\tilde{\epsilon}_l = s_l(n_l)t_{n_l-1, \alpha/2} / \sqrt{n_l}$.

In establishing a confidence interval on the optimality gap, $z^* - Eu_\lambda(\tilde{\xi}\hat{x})$, we must: (i) select \hat{x} , (ii) estimate $Eu_\lambda(\tilde{\xi}\hat{x})$, (iii) estimate z^* or an upper bound on z^* , and (iv) assess the sampling errors associated with (ii) and (iii). Given \hat{x} and the ability to generate iid observations of $(\tilde{\xi}, \tilde{r}_1, \tilde{r}_2)$, (ii) is straightforward to carry out and we do so in step 4 with point estimate $\bar{L}(n_l)$. No optimization is required in this computation; we simply compare the random returns $\tilde{\xi}^i \hat{x}$ with the benchmarks \tilde{r}_1^i and \tilde{r}_2^i and form a sample mean estimate of $Eu_\lambda(\tilde{\xi}\hat{x})$. The associated sampling error is denoted $\tilde{\epsilon}_l$ in step 5, where $t_{n_l-1, \alpha/2}$ is a $(1 - \alpha/2)$ -level quantile of a t -distribution with $n_l - 1$ degrees of freedom. With respect to (iii), a point estimate of an upper bound on z^* is given by z_n^* because $Ez_n^* \geq z^*$ (see Mak et al., 1999). Estimates of the associated sampling error (denoted $\tilde{\epsilon}_u$ in step 5) are formed by generating multiple observations of z_n^* as is done in step 2. In carrying out step 2, we solve n_u instances of (7) and in so doing we obtain n_u potential candidate solutions $x_n^{*1}, \dots, x_n^{*n_u}$. We combine these candidates via an

average in step 3 to obtain \hat{x} . (Other possibilities exist but averaging worked well in our computational experiments.) The confidence interval width on the optimality gap as stated in (8) is $\tilde{\epsilon} = [\bar{U}(n_u) - \bar{L}(n_l)]^+ + \tilde{\epsilon}_u + \tilde{\epsilon}_l$.

In steps 1 and 4 of the solution and quality-testing procedure iid observations of $(\tilde{\xi}, \tilde{r}_1, \tilde{r}_2)$ must be generated. We do so using the NORTA method as described in Section 3 with continuous empirical marginals and historical correlation matrix. In terms of the solution procedure's input parameters, we select $\alpha = 0.05$ so that the approximate confidence interval holds at level 0.95. As suggested above, evaluation of $\bar{L}(n_l)$ is inexpensive and we use a relatively large sample size of $n_l = 100,000$ in Section 5. The bulk of the computational effort in carrying out the procedure is in executing step 2. We recommend that sample size n_u be fixed at $n_u = 30$. Solution quality tends to improve, and the confidence interval width shrink, as n grows. Often the choice of n is driven, at least in part, by what can be handled with reasonable computational effort, and in the next section we use $n = 500$.

5. Computation

We illustrate our approach by constructing a fund of funds by using hedge-fund index data from the CSFB (Credit Suisse First Boston)/Tremont Hedge Fund Index (see the site <http://www.hedgeindex.com>). CSFB/Tremont maintains monthly return data for the following $m = 13$ indices: convertible arbitrage, dedicated short bias, emerging markets, equity market neutral, event driven, distressed, ED multi-strategy, risk arbitrage, fixed income arbitrage, global macro, long/short, managed futures, and multi-strategy. The returns cover the time period from April

Table 1
Summary statistics for CSFB/Tremont Hedge Fund indices for the period April 1994–March 2003

Index	Return (%)	Best month (%)	Worst month (%)	Std. dev. (%)	Skewness (%)	Kurtosis
Convertible arbitrage	10.63	3.57	−4.68	4.85	−1.67	4.34
Dedicated short bias	0.60	22.71	−8.69	18.17	0.89	2.15
Emerging markets	3.88	16.42	−23.03	18.02	−0.55	3.61
Equity market neutral	11.04	3.26	−1.15	3.13	0.13	0.15
Event driven	10.31	3.58	−11.77	6.14	−3.53	22.89
Distressed	12.08	4.10	−12.45	7.11	−2.81	16.26
ED multi-strategy	9.39	4.66	−11.52	6.50	−2.66	16.18
Risk arbitrage	7.97	3.81	−6.15	4.58	−1.30	5.87
Fixed income arbitrage	6.71	2.02	−6.96	4.05	−3.55	19.16
Global macro	14.21	10.60	−11.55	12.55	−0.02	1.92
Long/short	11.25	13.01	−11.43	11.37	0.25	3.18
Managed futures	6.85	9.95	−9.35	12.32	0.03	0.51
Multi-strategy	8.95	3.61	−4.76	4.62	−1.28	3.19

“Return” is the average annualized return over 108 months, and the standard deviation of return, “Std. dev.” is also annualized. The extreme monthly return values are labeled “Best Month” and “Worst Month.”

1994 through March 2003, for a total of 108 monthly returns. Summary statistics are reported in [Table 1](#).

A brief examination of the reported statistics indicate that the hedge-fund return distributions are non-normal. In addition, most of the indices fail standard statistical tests for normality (conducted in *SPLUS*). For each index, we construct its continuous empirical distribution, as described at the end of [Section 3](#), by using the historical $N = 108$ monthly returns. We also computed the sample correlation matrix based on these returns and used it as the nominal input, Σ_{ξ} , for the *NORTA* procedure described in [Section 3](#).

Our expected utility combines the goals of maximizing the probability of outperforming 10% annual return and minimizing the expected loss. In our notation defining u_{λ} , we use $r_1 = 10\%$ and $r_2 = 0\%$. If a portfolio manager had only mean–variance tools and wished to incorporate these benchmarks, he could construct two portfolios, maximizing the Sharpe ratio with the risk-free rate replaced by 10% in the first portfolio and by 0% in the second. [Table 2](#) shows these two mean–variance (MV) portfolios. The maximum Sharpe ratio portfolio at 0% is almost the minimum variance portfolio. The highest allocation is in the Equity Market Neutral index with the Fixed Income Arbitrage and Risk Arbitrage indices forming the next largest shares. The maximum Sharpe ratio portfolio using the 10% benchmark splits the allocation between the Equity Market Neutral, Distressed Debt and Global Macro Strategies indices.

Next, we solved the problem by using the methodology developed in [Sections 2–4](#). [Fig. 1](#) plots the efficient frontier when the risk term is expected regret with respect to 0% and the reward is the probability of outperforming 10% per year, as we vary the

Table 2
Mean–variance optimal portfolios obtained by maximizing the Sharpe ratio with the risk-free rate replaced by 0% and 10%

Index	MV at 0% (%)	MV at 10% (%)
Convertible arbitrage	0	0
Dedicated short bias	8	0
Emerging markets	0	0
Equity market neutral	51	40
Event driven	0	0
Distressed	6	25
ED multi-strategy	0	0
Risk arbitrage	10	0
Fixed income arbitrage	12	0
Global macro	0	35
Long/short	4	0
Managed futures	0	0
Multi-strategy	8	0
Probability of outperforming 10%	45	57.5
Expected regret with respect to 0%	0.75	2.90
Mean	9.24	12.78
Volatility	2.09	5.62

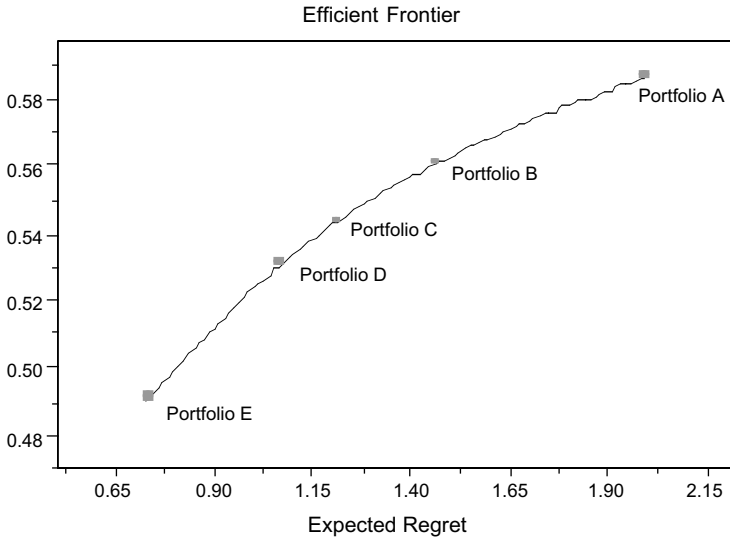


Fig. 1. Efficient frontier when the risk term is the expected regret with respect to 0% and the reward is the probability of outperforming 10%.

value of $\lambda \in \{0, 0.01, \dots, 0.09, 0.1, 0.2, \dots, 1.0\}$. Five specific portfolios are labeled on Fig. 1 and these are further detailed in Table 3. Each of these portfolios was formed by solving $n_u = 30$ instances of (7), each with $n = 500$ scenarios. The \hat{x} values shown in Table 3 were formed by averaging the 30 solutions obtained, as indicated in step 3 of the procedure in Section 4. The point estimates and confidence intervals of $P(\tilde{R}_{\hat{x}} \geq r_1)$ and $E(r_2 - \tilde{R}_{\hat{x}})^+$ shown in Table 3 were formed using $n_\ell = 100,000$ observations of ξ .

The plot in Fig. 1 indicates the efficient frontier appears concave. We emphasize that while similar forms were obtained in other data sets we have studied, this form is not obtained in general, due to the use of $P(\tilde{R}_x \geq r_1)$ in our expected utility. Specifically, $\phi(\ell)$ defined as

$$\begin{aligned} \phi(\ell) = \max_{x \in X, v, y} \quad & \sum_{\omega \in \Omega} p^\omega y^\omega, \\ \text{s.t.} \quad & \zeta^\omega x \geq r_1^\omega y^\omega - M^\omega (1 - y^\omega), \omega \in \Omega, \\ & v^\omega \geq r_2^\omega - \zeta^\omega x, \omega \in \Omega, \\ & \sum_{\omega \in \Omega} p^\omega v^\omega \leq \ell, \\ & y^\omega \in \{0, 1\}, v^\omega \geq 0, \omega \in \Omega, \end{aligned}$$

gives the efficient frontier as a function of the expected regret level ℓ when the underlying distribution is discrete. $\phi(\ell)$ is non-decreasing, piecewise constant and upper semicontinuous but is not, in general concave. (Moreover, neither the convex hull of the epi-graph of ϕ nor its complement is a convex set, in general.) This is in con-

Table 3

Efficient portfolios when the risk term is the expected regret with respect to 0% and the reward is the probability of outperforming 10%

Index	Portfolio A	Portfolio B	Portfolio C	Portfolio D	Portfolio E
Convertible arbitrage	53%	41%	35%	25%	9%
Dedicated short bias	2%	2%	2%	2%	2%
Emerging markets	1%	1%	1%	0%	0%
Equity market neutral	14%	23%	29%	36%	45%
Event driven	4%	5%	4%	5%	6%
Distressed	8%	7%	7%	6%	5%
ED multi-strategy	2%	2%	2%	1%	1%
Risk arbitrage	2%	2%	2%	3%	7%
Fixed income arbitrage	1%	3%	4%	3%	5%
Global macro	8%	7%	6%	5%	4%
Long/short	3%	2%	2%	2%	2%
Managed futures	1%	2%	2%	3%	4%
Multi-strategy	3%	4%	6%	8%	8%
Probability	58.3% (0.15)	56.5% (0.15)	55% (0.15)	53.1% (0.15)	50% (0.15)
Expected regret	1.99% (0.25)	1.56% (0.16)	1.32% (0.12)	1.03% (0.080)	0.78% (0.049)
Mean	10.75%	10.61%	10.42%	10.33%	10.07%
Volatility	3.98%	3.56%	3.25%	2.96%	2.73%

The reported probability is computed by using 10% as a benchmark and the expected regret uses 0% as a benchmark. The values in parentheses are half-widths for 95% confidence intervals associated with these point estimates using a sample size of 100,000.

trast to the Markowitz mean–variance setting in which the analogous $\phi(\ell)$, defined by maximizing $E\tilde{R}_x$ subject to $\text{Var}\tilde{R}_x \leq \ell$, is concave.

In a computational study in Morton et al. (2003) we describe a number of enhancements which significantly improve our ability to solve instances of (7). With respect to parameter $\lambda \geq 0$, the most computationally difficult version of (7) is that with $\lambda = 0$ in which the objective is to outperform benchmark r_1 with maximum probability. In Morton et al. (2003) we can solve 30 instances of (7) with $\lambda = 0$ and $n = 500$ scenarios in roughly three minutes on a 1.6 GHz Pentium IV computer with 512 Mb of memory. The computational effort required to solve the problems reported here was similar.

Portfolio A achieves benchmark r_1 with the highest probability, 58.3%. Portfolios B–E are optimal for the combined objective as the value of λ increases. Portfolio E essentially plays the role of the minimum-risk portfolio, similar to the minimum-variance portfolio in mean–variance analysis. Note that the main change in the allocation, as the risk aversion parameter increases, is that weight is moved from the

Table 4

Comparison between the portfolio obtained by maximizing the probability of outperforming 10% return (Portfolio A) and the mean–variance portfolio obtained by maximizing the Sharpe ratio using 10% instead of the risk-free rate

	Portfolio A (%)	Max Sharpe ratio at 10% portfolio (%)
Probability	58.3	57.5
Expected regret	1.99	2.90
Mean	10.75	12.78
Volatility	3.98	5.62

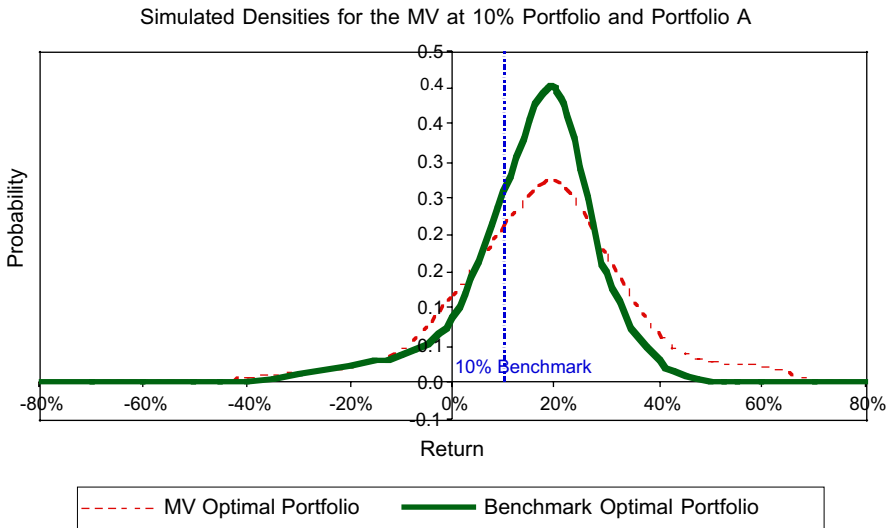


Fig. 2. Estimated densities for the mean–variance optimal portfolio (obtained by maximizing the Sharpe ratio at 10%) and Portfolio A (obtained by maximizing the probability of outperforming 10%).

convertible arbitrage index to the Equity Market Neutral index. By doing this, the probability of outperforming r_1 and the expected regret relative to r_2 both reduce in value.

One reasonable question to consider is what is the difference between the 10% Sharpe-ratio portfolio and the portfolio obtained by maximizing the probability of beating this benchmark (i.e., Portfolio A). In Table 4 we compare their performance with respect to: the probability of outperforming r_1 , expected regret relative to r_2 , mean return, and volatility. Portfolio A is dominant in all the measures except mean return. It achieves the goal of maximizing the probability of outperforming 10% by investing in a portfolio that has 40% smaller volatility than the mean–variance optimal portfolio. Additionally, the expected regret with respect to 0% is 53% smaller than the expected regret of the mean–variance optimal portfolio.

To continue the comparison, we used the NORTA procedure to simulate 10,000 observations of the return vector and estimated the corresponding densities of \tilde{R}_x when x is given by Portfolio A and by the mean–variance 10% Sharpe-ratio portfo-

lio. Fig. 2 displays the estimated densities. The mean–variance portfolio tries to achieve the benchmark of 10% by using assets with higher volatility, with 35% allocated to the Global Macro Strategy index that has 12.55% annual volatility and 25% to the Distressed Debt index that has 7.11% annual volatility. As a result the probability of outperformance is almost the same as Portfolio A but the price is a portfolio with higher variance. In contrast, Portfolio A achieves the desired benchmark by investing in assets that could provide returns close to the 10% benchmark. The resulting portfolio has an allocation of 53% in the convertible arbitrage index, which has 4.85% annual volatility, yielding a lower-variance portfolio relative to the mean–variance optimal portfolio.

6. Summary

The mean–variance approach to portfolio allocation has been the standard industry methodology for fifty years. The approach is founded on the assumption of normal distributions and an objective function of maximizing risk-adjusted return.

In this paper, we introduce a more general and flexible framework for asset allocation. Our approach utilizes recent advances in stochastic programming and in modeling, and simulating observations from, a multi-variate distribution. The approach is able to recognize and use the information embedded in the unusual return distributions of hedge funds. Importantly, the model uses a utility function that explicitly recognizes benchmarks, allowing for a broad range of preferences that are intuitively closer to the real objectives of investors.

To illustrate the method's ability to work with non-normal distributions, we construct a fund of hedge funds. The proposed framework can be used to construct a fund of funds particularly tailored to the investors' objectives. It can also be used to actively manage a portfolio of hedge funds with the goal of systematically maintaining performance above a given benchmark.

The need for an approximation method, like the one described in Section 4, arises in asset allocation models when all but the simplest of distributional assumptions and utility functions are employed. Our Monte Carlo approximation follows quite naturally from the use of the NORTA method (Section 3) for generating non-normal return variates. We emphasize that while our approach addresses the non-standard case of asset allocation for hedge-fund instruments, it also applies directly to other asset allocation problems in which handling benchmarks and non-normal return distributions are critical.

Acknowledgements

This research was sponsored by Deutsche Asset Management, Deutsche Bank, New York. Much of the work was done while Ivilina Popova was a Director, and David Morton and Elmira Popova were consultants, at the Global Research Center of Deutsche Asset Management, New York. David Morton's work was also partially

supported by the National Science Foundation under Grants DMI-9702217 and DMI-0217927.

References

- Agarwal, V., Naik, N.Y., 2001. Risks and portfolio decisions involving hedge funds. Technical Report, London Business School.
- Artzner, P., Delbaen, F., Eber, J.-M., Heath, D., 1997. Thinking coherently. *Risk* 10, 68–71.
- Artzner, P., Delbaen, F., Eber, J.-M., Heath, D., 1999. Coherent measures of risk. *Mathematical Finance* 9, 203–228.
- Attouch, H., Wets, R.J.-B., 1990. Epigraphical processes: Laws of large numbers for random lsc functions. Séminaire d'Analyse Convexe, Montpellier.
- Brooks, C., Kat, H.M., 2002. The statistical properties of hedge fund index returns and their implications for investors. *Journal of Alternative Investments* 5, 26–44.
- Browne, S., 1999. Beating a moving target: Optimal portfolio strategies for outperforming a stochastic benchmark. *Finance and Stochastics* 3, 275–294.
- Cario, M.C., Nelson, B.L., 1997. Modeling and generating random vectors with arbitrary marginal distributions and correlation matrix. Technical Report, Department of Industrial Engineering and Management Sciences, Northwestern University.
- Dembo, R.S., King, A.J., 1992. Tracking models and the optimal regret distribution in asset allocation. *Applied Stochastic Models and Data Analysis* 8, 151–157.
- Fishman, G.S., 1996. Monte Carlo: Concepts, Algorithms, and Applications. Springer Series in Operations Research. Springer-Verlag, New York.
- Fung, W., Hsieh, D.A., 1997. Empirical characteristics of dynamic trading strategies: The case of hedge funds. *The Review of Financial Studies* 10, 275–302.
- Fung, W., Hsieh, D.A., 2001. The risk in hedge fund strategies: Theory and evidence from trend followers. *The Review of Financial Studies* 14, 313–341.
- Ghosh, S., Henderson, S.G., 2002. Chessboard distributions and random vectors with specified marginals and covariance matrix. *Operations Research* 50, 820–834.
- Ghosh, S., Henderson, S.G., 2003. Behaviour of the NORTA method for correlated random vector generation as the dimension increases. *ACM Transactions on Modeling and Computer Simulation* 13, 276–294.
- King, A.J., Jensen, D.L., 1992. Linear-quadratic efficient frontiers for portfolio optimization. *Applied Stochastic Models and Data Analysis* 8, 195–207.
- Law, A.M., Kelton, W.D., 2000. *Simulation Modeling and Analysis*, third ed. McGraw-Hill, Boston.
- Mak, W.K., Morton, D.P., Wood, R.K., 1999. Monte Carlo bounding techniques for determining solution quality in stochastic programs. *Operations Research Letters* 24, 47–56.
- Markowitz, H., 1952. Portfolio selection. *Journal of Finance* 7, 77–91.
- Martin, G., 2001. Making sense of hedge fund returns: A new approach. In: Acar, E. (Ed.), *Added Value in Financial Institutions: Risk or Return?* Prentice Hall, Bet, pp. 165–182.
- Morton, D.P., Popova, E., Popova, I., Zhong, M., 2003. Optimizing benchmark-based utility functions. *Bulletin of the Czech Econometric Society* 10, 1–18.
- Morton, D.P., Wood, R.K., 1998. On a stochastic knapsack problem and generalizations. In: Woodruff, D.L. (Ed.), *Advances in Computational and Stochastic Optimization, Logic Programming, and Heuristic Search: Interfaces in Computer Science and Operations Research*. Kluwer Academic Publishers, Boston, pp. 149–168.
- Popova, I., Morton, D.P., Popova, E., 2003. Optimal hedge fund allocation with asymmetric preferences and distributions. Technical Report, The University of Texas at Austin.
- Rockafellar, R.T., Uryasev, S., 2000. Optimization of conditional value-at-risk. *Journal of Risk* 2, 21–41.
- Rockafellar, R.T., Uryasev, S., 2002. Conditional value-at-risk for general loss distributions. *Journal of Banking and Finance* 26, 1443–1471.