Proceedings of the 2007 Winter Simulation Conference
S. G. Henderson, B. Biller, M.-H. Hsieh, J. Shortle, J. D. Tew, and R. R. Barton, eds.

A CONFIDENCE INTERVAL FOR TAIL CONDITIONAL EXPECTATION VIA TWO-LEVEL SIMULATION

Hai Lan
Barry L. Nelson
Jeremy Staum

Department of Industrial Engineering and Management Sciences
Northwestern University
Evanston, IL 60208-3119, U.S.A.

ABSTRACT

We develop and evaluate a two-level simulation procedure that produces a confidence interval for tail conditional expectation, otherwise known as conditional tail expectation. This risk measure is closely related to conditional value-at-risk, expected shortfall, and worst conditional expectation. The outer level of simulation generates risk factors and the inner level estimates each expected loss conditional on the risk factor. Our procedure uses the statistical theory of empirical likelihood to construct a confidence interval, and it uses tools from the ranking-and-selection literature to make the simulation efficient.

1 INTRODUCTION

A fundamental task in risk management is to measure the risk entailed by a decision, such as the choice of a portfolio. Let V be a random variable representing the value resulting from the decision, and F_V be its distribution. A risk measure is a functional $T(F_V)$ of this distribution. For example, value at risk VaR_p may be defined as the negative of the p-quantile of F_V. In this paper, we are primarily interested in tail conditional expectation:

$$T(F_V) = \mathrm{TCE}_p := \mathrm{E}[-V | V \leq -\mathrm{VaR}_p].$$

Suppose that we wish to estimate $T(F_V)$ by simulation, but we are unable to sample from F_V. For example, let V be the gain experienced by a portfolio containing derivative securities. We may have a model of underlying financial markets that allows us to sample a risk factor Z (such as a stock price) from its distribution F_Z, yet be unable to evaluate the portfolio's gain $V := V(Z)$ because the function $V(\cdot)$ is unknown when the prices of derivative securities are unknown functions of the risk factor. However, we may be able to represent $V(Z) = \mathrm{E}[X|Z]$ where X involves the payoffs of derivative securities, which we can simulate conditional on the risk factor Z. In general, Z may be a vector. Then we can estimate the risk measure $T(F_V)$ by a two-level simulation in which the outer level of simulation generates k risk factors $Z_1, Z_2, \ldots, Z_k$ and the inner level estimates each $V_i := V(Z_i)$. For more on this framework and its significance in risk management, see Lan, Nelson, and Staum (2007). A subtlety in the financial context is that, while Z is sampled from its real-world probability distribution, X is a discounted payoff sampled from a risk-neutral distribution, which is required to make $V(Z) = \mathrm{E}[X|Z]$: see, for instance, Björk (1998).

In reality, risk management simulations may deal with thousands of complicated securities and non-trivial models. However, for purposes of illustration, we will consider the following example, which is also used in our experiments. At time 0, we sell a put option with strike price $K = 110$ and maturity $U = 1$ year on a stock whose initial price is $S_0 = 100$. This stock's price obeys the Black-Scholes model with drift $\mu = 6\%$ and $\sigma = 15\%$. There is a money market account with interest rate $r = 6\%$. The initial price for which we sell the put option is $P_0 = P(U, S_0)$, which is the Black-Scholes formula evaluated for maturity U and stock price S_0.

We are interested in TCE_p at the $p = 1\%$ level at time $T = 1/52$ years, or one week. The risk factor Z is a standard normal random variable that determines the stock price at time T:

$$S_T = S_0 \exp\left(\left(\mu - \frac{\sigma^2}{2}\right)T + \sigma\sqrt{T}Z\right).$$

The payoff at maturity U, discounted to time T, from selling the put for an initial price of P_0 is

$$X = e^{-r(U-T)}\left(P_0 e^{rU} - (K - S_U)^+\right),$$

where

$$S_U = S_T \exp\left(\left(r - \frac{\sigma^2}{2}\right)(U-T) + \sigma\sqrt{U-T}Z'\right)$$

and Z' is a standard normal random variable independent of Z.

In this simple example, we can actually find the value

$$V = \mathrm{E}[X|Z] = P_0 e^{rT} - P(U-T, S_T),$$

using the Black-Scholes formula evaluated for maturity $U-T$ and stock price S_T. Furthermore, it is clear that V is strictly decreasing in Z, so we can compute that $\mathrm{VaR}_{0.01} \approx 2.92$ by evaluating V at $Z = z_{0.01}$, the standard normal first percentile. By numerical integration, we can also compute $\mathrm{TCE}_{0.01} \approx 3.39$, which will help us to evaluate the performance of our procedure, which does not compute V by using the Black-Scholes formula, but rather estimates it using inner-level simulation of payoffs at maturity.

Our goal is efficient computation of a confidence interval for TCE_p whose coverage probability can be proved to be at least nominal asymptotically and which is good at realistic sample sizes. See Lan, Nelson, and Staum (2007) for a discussion of how to construct an asymptotically valid confidence interval for two-level simulations, in which there is an interaction between the statistical uncertainty generated by the outer and inner levels of simulation. Here we use empirical likelihood (Owen 2001) to handle uncertainty at the outer level and the central limit theorem to handle uncertainty at the inner level. Two-level simulations can be extremely computationally expensive: given the available computational budget, they may produce very wide confidence intervals. To produce a narrower confidence interval given a fixed computational budget, our procedure uses screening with common random numbers and restarting, and it allocates different numbers of payoffs to different risk factors. The dominant computational cost in these two-level simulations is in the generation of payoffs. We simplify the analysis by ignoring other computational costs and making the total number of payoffs, over all risk factors, equal to a computational budget C.

2 PROCEDURE

This section presents a fixed-budget two-level simulation procedure for generating a confidence interval for TCE_p. It is adaptive in the sense that it learns, from a first phase of simulation, a good number of risk factors k and good numbers of payoffs $N_1, N_2, \ldots, N_k$ to generate conditional on each risk factor. For ease of presentation, we first consider a procedure in which k is fixed and only $N_1, N_2, \ldots, N_k$ are learned. The adaptive procedure is built from this simpler fixed-k procedure: its first phase uses k_0 risk factors to learn a good value of k, and its second phase is the fixed-k procedure with this value of k.

At the inner level, both procedures have a first stage in which n_0 payoffs are generated for every risk factor, using common random numbers. After the first stage, screening eliminates risk factors which are not likely to belong to the left tail, and sample sizes $N_1, N_2, \ldots, N_k$ are chosen; the sample size N_i is 0 if risk factor Z_i has been screened out. The budget constraint used in choosing these sample sizes is $kn_0 + \sum_{i=1}^{k} N_i \le C$. The first-stage data are discarded, a process called "restarting." In the second stage, N_i payoffs are generated conditional on the risk factor Z_i for each $i = 1, 2, \ldots, k$ using independent sampling, and a confidence interval is formed.

In Lan, Nelson, and Staum (2007), we described a framework for two-level simulation that generates a two-sided confidence interval $[\hat{L}, \hat{U}]$ with confidence level $1-\alpha$ where α can be decomposed as $\alpha = \alpha_o + \alpha_i$, representing errors due to the outer and inner levels of simulation, respectively. Here we further decompose $\alpha_i = \alpha_s + \alpha_{lo} + \alpha_{hi}$, where α_s is error due to screening and α_{lo} and α_{hi} are errors respectively associated with lower and upper confidence limits for inner-level simulation.

The procedures rely on empirical likelihood to account for statistical uncertainty at the outer level, that is, for the fact that $V_1, V_2, \ldots, V_k$ is only a sample from the true distribution F_V of portfolio values at time horizon T. The construction of an outer-level confidence interval for TCE_p based on empirical likelihood is discussed by Baysal and Staum (2007). Here we review a few essential facts for understanding the operation of empirical likelihood in our two-level simulation procedures.

Empirical likelihood involves assigning a vector $\mathbf{w}$ of weights to the vector $\mathbf{V} = (V_1, V_2, \ldots, V_k)$ of portfolio values. This weight vector $\mathbf{w}$ must belong to the set $\mathscr{S}$ of vectors that satisfy

$$\begin{aligned} \mathbf{w} &\ge 0, \\ \sum_{i=1}^{k} w_i &= 1, \\ \exists l \ni \sum_{i=1}^{l} w_i &= p, \quad \text{and} \\ \prod_{i=1}^{k} w_i &\ge c k^{-k}, \end{aligned}$$

where c is a critical value derived from a chi-squared distribution. The number l is a function of $\mathbf{w}$. Because TCE_p involves an average over the left tail containing probability p, we also define a transformed weight vector $\mathbf{w}'$, another

function of $\mathbf{w}$:

$$w_i' := \begin{cases} -\frac{w_i}{p}, & i = 1,2,\ldots,l \\ 0, & \text{otherwise.} \end{cases}$$

Let π_V be a permuation of $\{1,2,\ldots,k\}$ such that $V_{\pi_{V1}} \leq V_{\pi_{V2}} \leq \cdots \leq V_{\pi_{Vk}}$. If the vector $\mathbf{V} = (V_1, V_2, \ldots, V_k)$ of true portfolio values were known, the natural point estimator of TCE_p would be $\sum_{i=1}^{k} w_i' V_{\pi_{Vi}}$.

The use of empirical likelihood influences the way screening is done in our procedures. Given $\mathbf{w}$, to estimate $\sum_{i=1}^{k} w_i' V_{\pi_{Vi}}$ via inner-level simulation, we seek to identify which l risk factors belong to the left tail, that is, have indices $\pi_{V1}, \pi_{V2}, \ldots, \pi_{Vl}$. We do not need to know the ranking of risk factors within the left tail, nor outside the left tail. However, the empirical likelihood method of constructing a confidence interval involves optimization over $\mathbf{w} \in \mathscr{S}$, and different weight vectors have different values of l. From the properties of the set $\mathscr{S}$, we can deduce that all $\mathbf{w} \in \mathscr{S}$ have $l \leq l_{\max}$, where

$$l_{\max} := \max\left\{ l : k^k \left(\frac{p}{l}\right)^l \left(\frac{1-p}{k-l}\right)^{(k-l)} \geq c \right\}.$$

Screening eliminates all risk factors except those in a set I, which we want to contain $\{\pi_{V1}, \pi_{V2}, \ldots, \pi_{Vl_{\max}}\}$ with probability at least $1-\alpha_s$. We wish to eliminate those risk factors such that at least $l_{\max}$ other risk factors yield larger losses. The number of pairwise comparisons between $\{\pi_{V1}, \pi_{V2}, \ldots, \pi_{Vl_{\max}}\}$ and all other risk factors is $l_{\max}(k - l_{\max})$. Therefore, for each ordered pair (i, j) we do a hypothesis test that $V_i \geq V_j$ at level $\alpha_s/((k-l_{\max})l_{\max})$. If the test is true, then we say Z_i is "beaten" by Z_j. For each $i = 1,2,\ldots,k$, let $X_{i1}, X_{i2}, \ldots, X_{in_0}$ be an i.i.d. sample drawn from the conditional distribution of X given Z_i, and let $\bar{X}_i(n_0)$ be its sample average. For each $i, j = 1,2,\ldots,k$, let $S_{ij}^2(n_0)$ be the sample variance of $X_{i1} - X_{j1}, X_{i2} - X_{j2}, \ldots, X_{in_0} - X_{jn_0}$. We screen out all risk factors which are beaten at least $l_{\max}$ times:

$$I = \left\{ i : \sum_{i \neq j} \mathbf{1}\left\{ \bar{X}_i(n_0) > \bar{X}_j(n_0) + d\frac{S_{ij}(n_0)}{\sqrt{n_0}} \right\} < l_{\max} \right\}$$

where $\mathbf{1}\{\cdot\}$ is an indicator function and

$$d = t_{n_0-1, 1-\alpha_s/((k-l_{\max})l_{\max})}$$

is the $1-\alpha_s/((k-l_{\max})l_{\max})$ quantile of the t-distribution with $n_0 - 1$ degrees of freedom. Computing I could require as many as $k(k-1)/2$ comparisons. However, it is generally unnecessary to do all of them: it is possible to screen out many risk factors with only one comparison each. The procedure in the next section is set up to take advantage of this opportunity.

2.1 A Procedure with Fixed Number of Risk Factors k

The procedure has the following steps:

1. **Risk Factor Generation**:
   Generate risk factors $Z_1, Z_2, \ldots, Z_k$ independently from the distribution F_Z.
2. **First Stage Sampling**:
   Sample payoffs $X_{ij}, j = 1,2,\ldots,n_0$, conditionally on Z_i with common random numbers, for each $i = 1,2,\ldots,k$.
3. **Screening**:
   Initialize $I \leftarrow \{1,2,\ldots,k\}$.
   For each $i = 1,2,\ldots,k$, compute the sample average $\bar{X}_i(n_0)$ of $X_{i1}, X_{i2}, \ldots, X_{in_0}$.
   Sort to produce a permutation π_0 of $\{1,2,\ldots,k\}$ such that $\bar{X}_{\pi_{0i}}(n_0)$ is nondecreasing in i.
   For each $i = k, k-1, \ldots, l_{\max}+1$, do:
   (a) Initialize $j \leftarrow 1$, $b \leftarrow 0$, and $g \leftarrow 1$.
   (b) Do while $j \leq k$ and $g = 1$:
      i. Compute the sample variance $S^2_{\pi_{0i}\pi_{0j}}(n_0)$ of $X_{\pi_{0i}1} - X_{\pi_{0j}1}, X_{\pi_{0i}2} - X_{\pi_{0j}2}, \ldots, X_{\pi_{0i}n_0} - X_{\pi_{0j}n_0}$.
      ii. Set $b \leftarrow b + \mathbf{1}\{\bar{X}_{\pi_{0i}}(n_0) > \bar{X}_{\pi_{0j}}(n_0) + dS_{\pi_{0i}\pi_{0j}}(n_0)/\sqrt{n_0}\}$.
      iii. If $b \geq l_{\max}$, set $I \leftarrow I \setminus \{i\}$ and $g \leftarrow 0$.
      iv. Set $j \leftarrow j+1$.
4. **Restarting and Second Stage Sampling**:
   For each $i \in I$, compute the sample variance $S_i^2(n_0)$ of $X_{i1}, X_{i2}, \ldots, X_{in_0}$.
   Discard all payoffs from Step 2.
   For each $i \in I$, set

$$N_i \leftarrow \frac{(C - kn_0)S_i^2(n_0)}{\sum_{j \in I} S_j^2(n_0)}.$$

   Sample payoffs $X_{ij}, j = 1,2,\ldots,N_i$, conditionally on Z_i and independently for each $i \in I$.
5. **Constructing the Confidence Interval**:
   For each $i \in I$, compute the sample average $\bar{X}_i(N_i)$ and sample variance $S_i^2(N_i)$ of $X_{i1}, X_{i2}, \ldots, X_{iN_i}$.
   Sort to produce a permutation π_L of I such that $\bar{X}_{\pi_{Li}}(N_{\pi_{Li}}) + z_{lo}S_{\pi_{Li}}(N_{\pi_{Li}})/\sqrt{N_{\pi_{Li}}}$ is nondecreasing in i, where z_{lo} is the $(1-\alpha_{lo})^{1/|I|}$ quantile of the standard normal distribution.

The lower confidence limit is

$$\hat{L} = \min_{\mathbf{w}\in\mathscr{S}} \sum_{i=1}^{l_{\max}} w_i' \left(\bar{X}_{\pi_{Li}}(N_{\pi_{Li}}) + z_{lo} \frac{S_{\pi_{Li}}(N_{\pi_{Li}})}{\sqrt{N_{\pi_{Li}}}} \right). \tag{1}$$

Sort to produce a permutation π_1 of I such that $\bar{X}_{\pi_{1i}}(N_{\pi_{1i}})$ is nondecreasing in i.
Sort to produce a permutation π_S of I such that $S^2_{\pi_{Si}}(N_{\pi_{Si}})/N_{\pi_{Si}}$is nonincreasing in i.
The upper confidence limit is

$$\hat{U} = \max_{\mathbf{w}\in\mathscr{S}} \sum_{i=1}^{l_{\max}} w_i' \bar{X}_{\pi_{1i}}(N_{\pi_{1i}}) + z_{hi} B \tag{2}$$

where z_{hi} is the $1-\alpha_{hi}$ quantile of the standard normal distribution and

$$B := \max_{\mathbf{w}\in\mathscr{S}} \sqrt{\sum_{i=1}^{l_{\max}} (w_i')^2 \frac{S^2_{\pi_{Si}}(N_{\pi_{Si}})}{N_{\pi_{Si}}}}$$

is computed by using the Newton method to solve its KKT conditions (see Bazaraa et al. 2005).

An algorithm for the optimizations in Equations (1)–(2) is given in Baysal and Staum (2007).

2.2 An Adaptive Procedure

In the first phase of this procedure, k_0 risk factors are generated. The first phase collects information used to choose the number of risk factors k for the second phase. In the following, we append k_0 or k, or n_0 or N_i, to quantities computed on the basis of k_0 or k risk factors, or from n_0 or N_i payoffs. The choice of k is based on the following conjectures:

1. The width of the outer-level empirical likelihood confidence interval with k risk factors will be approximately $E_o(k_0)/\sqrt{k}$, where

$$\begin{aligned} E_o(k_0) &= \sqrt{k_0} \left(\max_{\mathbf{w}\in\mathscr{S}(k_0)} \sum_{i=1}^{k_0} w_i' \bar{X}_{\pi_{0i}}(n_0) \right. \\ &\quad \left. - \min_{\mathbf{w}\in\mathscr{S}(k_0)} \sum_{i=1}^{k_0} w_i' \bar{X}_{\pi_{0i}}(n_0) \right), \end{aligned}$$

   so that $E_o(k_0)/\sqrt{k_0}$ is the width of the outer-level empirical likelihood confidence interval with k_0 risk factors.
2. The fraction of risk factors that survive screening does not depend on the number of risk factors: $|I(k)|/k \approx |I(k_0)|/k_0$.
3. The average variance of the payoff conditional on risk factors that survive screening does not depend on the number of risk factors: $\sum_{i\in I(k)} S_i^2(n_0)/|I(k)| \approx \sum_{i\in I(k_0)} S_i^2(n_0)/|I(k_0)|$.
4. After the second stage, the sample variance will be approximately the same as it was after the first stage, $S_i^2(N_i) \approx S_i^2(n_0)$, and therefore the number of payoffs chosen for each risk factor that survives screening will be approximately proportional to the conditional standard deviation: $S_i(N_i)/\sqrt{N_i}$ will be approximately constant over i. Based on the choice of N_i and the above conjectures,

$$\frac{S_i(N_i)}{\sqrt{N_i}} \approx \sqrt{\frac{\frac{1}{k_0}\sum_{j\in I(k_0)} S_j^2(n_0)}{C/k - n_0}}.$$

Given this, we can guess the width of the two-level confidence interval with k risk factors. First,

$$\begin{aligned} \hat{L} &= \min_{\mathbf{w}\in\mathscr{S}(k)} \sum_{i=1}^{l_{\max}} w_i' \left(\bar{X}_{\pi_{1i}}(N_{\pi_{1i}}) + z_{lo} \frac{S_{\pi_{1i}}(N_{\pi_{1i}})}{\sqrt{N_{\pi_{1i}}}} \right) \\ &\approx \min_{\mathbf{w}\in\mathscr{S}(k)} \sum_{i=1}^{l_{\max}} w_i' \bar{X}_{\pi_{1i}}(N_{\pi_{1i}}) + \sqrt{\frac{\bar{z}_{lo}(k)\sum_{j\in I(k_0)} S_j^2(n_0)}{k_0(C/k - n_0)}} \end{aligned}$$

where

$$\bar{z}_{lo}(k) := z_{(1-\alpha_{lo})^{k_0/(k|I(k_0)|)}} \approx z_{(1-\alpha_{lo})^{1/|I(k)|}} = z_{lo}.$$

Next,

$$\begin{aligned} \hat{U} &= \max_{\mathbf{w}\in\mathscr{S}(k)} \sum_{i=1}^{l_{\max}} w_i' \bar{X}_{\pi_{1i}}(N_{\pi_{1i}}) \\ &\quad + z_{hi} \max_{\mathbf{w}\in\mathscr{S}(k)} \sqrt{\sum_{i=1}^{l_{\max}} (w_i')^2 \frac{S^2_{\pi_{Si}}(N_{\pi_{Si}})}{N_{\pi_{Si}}}} \\ &\approx \max_{\mathbf{w}\in\mathscr{S}(k)} \sum_{i=1}^{l_{\max}} w_i' \bar{X}_{\pi_{1i}}(N_{\pi_{1i}}) \\ &\quad + z_{hi} \sqrt{\frac{\frac{1}{k_0}\sum_{j\in I(k_0)} S_j^2(n_0)}{C/k - n_0}} \max_{\mathbf{w}\in\mathscr{S}(k)} \sqrt{\sum_{i=1}^{l_{\max}} (w_i')^2}. \end{aligned}$$

Using the assumption that the outer-level confidence width with k risk factors

$$\max_{\mathbf{w}\in\mathscr{S}} \sum_{i=1}^{l_{\max}} w_i' \bar{X}_{\pi_{1i}}(N_{\pi_{1i}}) - \min_{\mathbf{w}\in\mathscr{S}} \sum_{i=1}^{l_{\max}} w_i' \bar{X}_{\pi_{1i}}(N_{\pi_{1i}}) \approx \frac{E_o(k_0)}{\sqrt{k}},$$

we can approximate the overall width $\hat{U} - \hat{L}$ by

$$E_o(k_0)/\sqrt{k} \tag{3}$$

$$+ \sqrt{\frac{\sum_{j\in I(k_0)} S_j^2(n_0)}{k_0(C/k - n_0)}} \left\{ \bar{z}_{lo}(k) + z_{hi} \max_{\mathbf{w}\in\mathscr{S}(k)} \sqrt{\sum_{i=1}^{l_{\max}} (w_i')^2} \right\}.$$

We choose k to minimize Equation (3), using the golden section method (see Bazaraa et al. 2005) within the interval of $[\lceil 2/p \rceil, \lfloor C/n_0 \rfloor]$.

The adaptive procedure has the following steps:

1. **First Phase Risk Factor Generation**:
   Generate risk factors $Z_1, Z_2, \ldots, Z_{k_0}$ independently from the distribution F_Z.
2. **First Phase, First Stage Sampling**:
   Sample payoffs $X_{ij}, j = 1,2,\ldots n_0$, conditionally on Z_i with common random numbers, for each $i = 1,2,\ldots,k_0$.
3. **Choose Number of Risk Factors**:
   Compute $I(k_0)$ as in Step 3 of Section 2.1, but do *not* eliminate the risk factors that are not in $I(k_0)$. Compute k that minimizes Equation (3).
4. **Second Phase Risk Factor Generation**:
   If $k > k_0$, generate risk factors $Z_{k_0+1}, Z_{k_0+2}, \ldots, Z_k$ independently from the distribution F_Z.
   If $k < k_0$, eliminate risk factors $Z_{k+1}, Z_{k+2}, \ldots, Z_{k_0}$.
5. **Second Phase, First Stage Sampling**:
   Sample payoffs $X_{ij}, j = 1,2,\ldots,n_0$, conditionally on Z_i with common random numbers, for each $i = k_0+1, k_0+2\ldots,k$.
6. **Second Stage**:
   Perform Steps 3–5 of Section 2.1.

3 EXPERIMENTS

We tested our procedures on the problem of producing a 90% confidence interval for $\text{TCE}_{0.01}$ in the example of the option described in Section 1. The number of payoffs generated for each risk factor in the first stage was $n_0 = 80$ in our two-stage procedures. The error $\alpha = 10\%$ was decomposed into $\alpha_o = 5\%$, $\alpha_s = 1\%$, $\alpha_{lo} = 2.5\%$, and $\alpha_{hi} = 1.5\%$.

For purposes of comparison, we also tested two other procedures. The four procedures we tested are:

1. The `efficient` fixed-k procedure described in Section 2.1.
2. The `adaptive` procedure of Section 2.2.
3. The `plain` procedure, which is similar to the efficient procedure, except that it does not use any efficiency techniques. It is a one-stage procedure that does not use screening (and therefore uses only independent sampling, not common random numbers), and it samples C/k payoffs for each risk factor.
4. The `rudimentary` procedure uses the same sampling scheme as the plain procedure, and then forms a confidence interval without taking any account of inner-level statistical uncertainty. It treats the vector of sample averages $\bar{\mathbf{X}}$ as if it were the vector of true values $\mathbf{V}$, and constructs the confidence interval described by Baysal and Staum (2007) based on the work of Manistre and Hancock (2005).

In Figure 1, we show the fraction of confidence intervals that covered the true value of $\text{TCE}_{0.01}$ in 100 independent macroreplications. Noticeable uncertainty about the coverage probability remains, so the figure also shows 95% confidence limits for each coverage probability. We see clearly that the rudimentary procedure yields extremely poor coverage: one can not produce a meaningful confidence interval while ignoring the statistical uncertainty generated by one of the two levels of the simulation experiment. The figure also shows that, at least for $k \geq 1{,}000$ risk factors, our procedures give at least nominal coverage. The coverage of the outer-level confidence level based on empirical likelihood, even in the absence of inner-level uncertainty, is also less than nominal for $k \leq 500$ risk factors in experiments reported by Baysal and Staum (2007). With so few risk factors, we do not know of a nonparametric method for generating a confidence interval for TCE_p with adequate coverage. With a large number of risk factors, our procedures yield excessive coverage. This is similar to results found by Lesnevski, Nelson, and Staum (2006): combatting the natural bias of the problem with ranking-and-selection tools, involving conservative probability inequalities, seems to cause wide confidence intervals with excessive coverage.

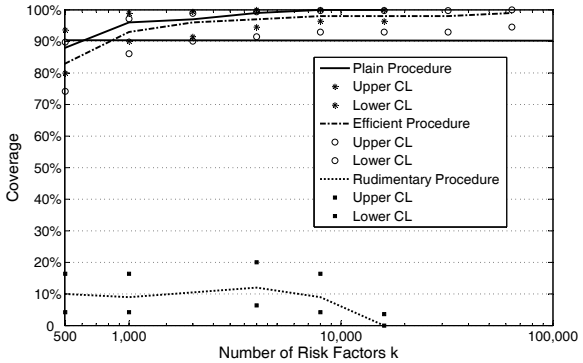

Figure 1: Coverage of fixed-k procedures with budget $C =$ 16 million.

Figure 2 illustrates the effect of efficiency techniques by comparing the plain and efficient fixed-k procedures with computational budget $C = 16$ million payoffs. The picture was similar for other values of C. The confidence interval width attainable for the best choice of k is about three times smaller for the efficient procedure than for the plain procedure, corresponding to a variance reduction factor of 9. The efficient procedure is best with about $k = 16{,}000$ risk factors for this budget, much more than the best value of k

for the plain procedure. The efficient procedure can afford to have more risk factors because it screens out many of them. For $k = 16{,}000$, the plain procedure can only give 1,000 payoffs to each risk factor, resulting in large inner-level uncertainty and confidence interval width so wide that it is off the chart. We see that there is a fairly wide range of k for which each procedure performs well, but an inappropriate choice of the number of risk factors k given the budget C can results in excessively wide confidence intervals. This motivated the development of the adaptive procedure.

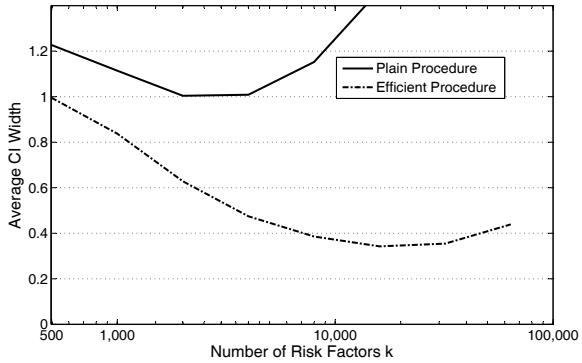

Figure 2: Confidence interval width of fixed-k procedures with budget $C = 16$ million.

Figure 3 illustrates that the adaptive procedure yields a confidence interval whose width is very close to that of the efficient fixed-k procedure *with the best choice of k*, the number of risk factors. This shows that the adaptive procedure, using $k_0 = 4{,}000$ risk factors initially, is able to identify very well the total number of risk factors to use, for the problem and range of budgets considered here. From Figure 2, we saw that the best k for the efficient fixed-k procedure is over 10,000 for budget $C = 16$ million. The coverage of the adaptive procedure is the same as that of the fixed procedure with the same choice of k. Because Figure 1 shows that the coverage of the fixed procedure was adequate with $C = 16$ million and $k = 10{,}000$, we conclude that the coverage of the adaptive procedure is adequate for $C = 16$ million. Analogous figures not published here support the same conclusions for the larger budgets we tested.

The curves on the log-log plot in Figure 3 are nearly linear with slope -0.25. This is unfavorable compared not only with the usual $\mathcal{O}(C^{-1/2})$ order of convergence of ordinary Monte Carlo, but even with the $\mathcal{O}(C^{-1/3})$ order of convergence for a two-level simulation estimator of VaR_p found by Lee (1998). Experiments with budgets from $C = 1$ million to 128 million revealed that the number of risk factors k chosen by the adaptive procedure increases roughly as $\mathcal{O}(C^{1/2})$. The outer-level width in Equation (3) was thus $\mathcal{O}(k^{-1/2}) = \mathcal{O}(C^{-1/4})$. We observed that the dominant term in the inner-level width in Equation (3) was the one due to the lower confidence limit, and its behavior was also similar to $\mathcal{O}(C^{-1/4})$.

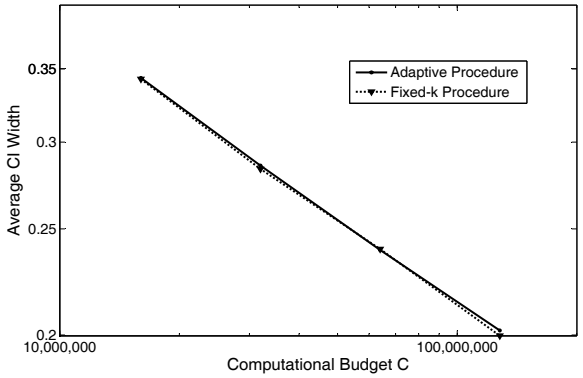

Figure 3: Confidence interval width of adaptive procedure vs. best fixed-k procedure.

4 PROOF

This section presents the elements of a proof that the fixed-k procedure of Section 2.1 yields a confidence interval with asymptotic coverage at least $1-\alpha$ as $k \to \infty$. We use the assumption that each payoff X_{ij} is normally distributed. In reality and in the example on which our experiments are run, payoffs are not normally distributed. However, sample averages of payoffs are approximately normal because of the central limit theorem if the sample sizes n_0 and N_i are large enough. Lesnevski, Nelson, and Staum (2006) studied the coverage of the confidence interval produced by a related procedure and found that a sample size of 30 was large enough to provide adequate coverage unless the payoffs are extremely far from normal. At the end of Section 4.2, we also make use of an approximation that is good when $N := \min_{i \in I} N_i$ is large.

We are working within the framework of Lan, Nelson, and Staum (2007): the two-level simulation confidence limits in Equations (1)–(2) arise from an inner-level confidence region $\mathscr{V}$ for the vector of true values $\mathbf{V} = (V_1, V_2, \ldots, V_k)$ as

$$\hat{L} = \min_{\mathbf{v} \in \mathscr{V}, \mathbf{w} \in \mathscr{S}} \sum_{i=1}^{k} w_i' v_{\pi_{vi}} \quad \text{and} \quad \hat{U} = \max_{\mathbf{v} \in \mathscr{V}, \mathbf{w} \in \mathscr{S}} \sum_{i=1}^{k} w_i' v_{\pi_{vi}}. \quad (4)$$

For any k-vector $\mathbf{v}$, let π_v be a permutation of $\{1, 2, \ldots, k\}$ such that $v_{\pi_{vi}}$ is nondecreasing in i. Our confidence region $\mathscr{V}$ for $\mathbf{V}$ is defined as the set containing all vectors $\mathbf{v}$ such that

$$\forall i \notin I,\ v_i \geq v_{\pi_{v l_{\max}}}, \quad (5)$$

$$\forall i \in I,\ v_i \leq \bar{X}_i(N_i) + z_{lo} S_i(N_i)/\sqrt{N_i}, \quad \text{and} \quad (6)$$

$$\max_{\mathbf{w} \in \mathscr{S}} \sum_{i=1}^{k} w_i' v_{\pi_{vi}} \leq \max_{\mathbf{w} \in \mathscr{S}} \sum_{i=1}^{l_{\max}} w_i' \bar{X}_{\pi_{1i}}(N_i) + z_{hi} B. \quad (7)$$

Notice that it does not matter whether we take the summations in Equation (7) from 1 to $l_{\max}$ or to $|I|$ or to k, because $l_{\max} \le |I| \le k$, and $w'_i = 0$ if $i > l_{\max}$ and $\mathbf{w} \in \mathscr{S}$. The notation merely emphasizes that π_{1i} does not exist for $i > |I| \ge l_{\max}$, because only $|I|$ risk factors survived screening to reach the second stage. Next we verify that Equations (1)–(2) follow from Equations (4)–(7).

First consider $\hat{L}$. Notice that $\sum_{i=1}^{k} w'_i v_{\pi_{vi}}$ is nonincreasing in $\mathbf{v}$. The minimum over $\mathbf{v}$ in Equation (4) is attained by taking each v_i to be as large as possible. Equation (7) does not provide a binding constraint. Take $v_i = +\infty$ for $i \notin I$ so that Equation (5) is satisfied and $\sum_{i=1}^{k} w'_i v_{\pi_{vi}}$ only depends on v_i for $i \in I$. For $i \in I$, take $v_i = \bar{X}_i(N_i) + z_{lo} S_i(N_i)/\sqrt{N_i}$. Using Equation (6) then shows that $\hat{L}$ as defined in Equation (4) satisfies Equation (1).

Next consider $\hat{U}$. Because of Equation (7), Equation (2) is an upper bound for the maximum in Equation (4). This upper bound is attained by taking $v_i = +\infty$ for $i \notin I$ and $v_i = \bar{X}_i(N_i) - z_{hi} B$ for $i \in I$. This choice of $\mathbf{v}$ makes $\pi_v = \pi_1$ and satisfies Equations (5)–(7).

If we plug in the true values $\mathbf{V}$ for $\mathbf{v}$ in the definition of $\mathscr{V}$, then the constraints defining $\mathscr{V}$ have the following interpretations.

- Equation (5) is equivalent to correct screening: $\gamma := \{\pi_{V1}, \pi_{V2}, \ldots, \pi_{Vl_{\max}}\} \subseteq I$.
- Equation (6) implies that $\hat{L} \le \min_{\mathbf{w} \in \mathscr{S}} \sum_{i=1}^{k} w'_i V_{\pi_{Vi}}$, the outer-level lower confidence limit. Because each π_{Li} is in I, Equations (1) and (6) imply that $\hat{L} \le \min_{\mathbf{w} \in \mathscr{S}} \sum_{i=1}^{l_{\max}} w'_i V_{\pi_{Li}}$. This is bounded above by $\min_{\mathbf{w} \in \mathscr{S}} \sum_{i=1}^{k} w'_i V_{\pi_{Vi}}$ because $\mathscr{S}$ is symmetric, $\mathbf{w}' \le 0$, and $\{V_{\pi_{V1}}, V_{\pi_{V2}}, \ldots, V_{\pi_{Vl_{\max}}}\}$ contains the lowest components of $\mathbf{V}$.
- Equation (7) implies that $\hat{U} \ge \max_{\mathbf{w} \in \mathscr{S}} \sum_{i=1}^{k} w'_i V_{\pi_{Vi}}$, the outer-level upper confidence limit.

Baysal and Staum (2007) show that the asymptotic probability that TCE_p is contained in the outer-level confidence interval is at least $1 - \alpha_o$ as $k \to \infty$. By the results of Lan, Nelson, and Staum (2007), it then suffices to show that the probability that $\mathbf{V} \in \mathscr{V}$ is at least $1 - \alpha_i$. By the Bonferroni inequality, the probability that $\mathbf{V} \notin \mathscr{V}$ is bounded above by the sum of the probabilities

- that $\mathbf{V}$ does not satisfy Equation (5),
- that $\mathbf{V}$ does not satisfy Equation (6), and
- that $\mathbf{V}$ satisfies Equation (5) and does not satisfy Equation (7).

In Section 4.1 we will show that the probability that $\mathbf{V}$ does not satisfy Equation (5) is bounded above by α_s. In Section 4.2 we will show that the latter two probabilities are bounded above by α_{lo} and α_{hi} respectively. This completes the proof because $\alpha_i = \alpha_s + \alpha_{lo} + \alpha_{hi}$.

4.1 Screening

Here we show that the probability of correct screening $\Pr\{\gamma \subseteq I\} \ge 1 - \alpha_s$. Let

$$B_{ij} := \mathbf{1}\{\bar{X}_i(n_0) > \bar{X}_j(n_0) + d S_{ij}(n_0)/\sqrt{n_0}\}$$

be the indicator function which is 1 when Z_j beats Z_i. We have

$$\begin{aligned} \Pr\{\gamma \subseteq I\} &= \Pr\left\{\forall i \in \gamma, \sum_{j \ne i} B_{ij} < l_{\max}\right\} \\ &\ge \Pr\left\{\forall i \in \gamma, j \notin \gamma, B_{ij} = 0\right\} \\ &\ge 1 - \sum_{i \in \gamma} \sum_{j \notin \gamma} \Pr\{B_{ij} = 1\} \end{aligned}$$

by the Bonferroni inequality. For $i \in \gamma$ and $j \notin \gamma$, $V_i \le V_j$. Therefore each

$$\begin{aligned} \Pr\{B_{ij} = 1\} &= \Pr\left\{\frac{\bar{X}_i(n_0) - \bar{X}_j(n_0)}{S_{ij}(n_0)/\sqrt{n_0}} > d\right\} \\ &\le \alpha_s/(l_{\max}(k - l_{\max})), \end{aligned}$$

using $d = t_{n_0 - 1, 1 - \alpha_s/(k - l_{\max}) l_{\max}}$.

4.2 Confidence Region

In this section, we deal with the second-stage inner-level simulation, after screening and restarting have occurred. We can think of the first stage as randomly generating a simulation problem which the second stage solves. The first stage produces I and N_i for each $i \in I$. This is an experimental design for the second stage, specifying which risk factors to consider and how many payoffs to simulate from each of them.

The probability that $\mathbf{V}$ does not satisfy Equation (6) is

$$\begin{aligned} &\Pr\left\{\exists i \in I \ni V_i > \bar{X}_i(N_i) + \frac{S_i(N_i)}{\sqrt{N_i}} z_{lo}\right\} \\ &= 1 - \prod_{i \in I} \Pr\left\{V_i \le \bar{X}_i(N_i) + \frac{S_i(N_i)}{\sqrt{N_i}} z_{lo}\right\} \end{aligned}$$

because sampling is independent at the second stage. This is

$$1 - \left((1 - \alpha_{lo})^{1/|I|}\right)^{|I|} = \alpha_{lo}.$$

Finally, we consider the probability that $\mathbf{V}$ satisfies Equation (5) and does not satisfy Equation (7). The following argument will show that, conditional on any experimental design such that correct screening occurs after the first stage (that is, $\mathbf{V}$ satisfies Equation (5)), the probability that the second-stage simulation causes $\mathbf{V}$ not to satisfy Equation (7)

is bounded above by α_{hi}. Consequently, the probability that $\mathbf{V}$ satisfies Equation (5) and does not satisfy Equation (7) is bounded above by α_{hi}. Define

$$\mathbf{w}^* := \arg\max_{\mathbf{w}\in\mathscr{S}} \sum_{i=1}^{k} w_i' V_{\pi_{Vi}}.$$

Then

$$\max_{\mathbf{w}\in\mathscr{S}} \sum_{i=1}^{l_{\max}} w_i' \bar{X}_{\pi_{1i}}(N_{\pi_{1i}}) \geq \sum_{i=1}^{l_{\max}} w_i^{*\prime} \bar{X}_{\pi_{1i}}(N_{\pi_{1i}}).$$

By the optimality properties shown in Baysal and Staum (2007), w_i^* is nonincreasing in i, so $w_i^{*\prime}$ is nondecreasing. Because $\bar{X}_{\pi_{1i}}(N_{\pi_{1i}})$ is nondecreasing in i,

$$\sum_{i=1}^{l_{\max}} w_i^{*\prime} \bar{X}_{\pi_{1i}}(N_{\pi_{1i}}) \geq \sum_{i=1}^{l_{\max}} w_i^{*\prime} \bar{X}_{\pi_i}(N_{\pi_i})$$

for any permutation π of I. Because correct screening occurred, that is, $\gamma := \{\pi_{V1}, \pi_{V2}, \ldots, \pi_{Vl_{\max}}\} \subseteq I$, there exists a permutation π of I such that $\pi_i = \pi_{Vi}$ for $i = 1, 2, \ldots, l_{\max}$. Consequently,

$$\max_{\mathbf{w}\in\mathscr{S}} \sum_{i=1}^{l_{\max}} w_i' \bar{X}_{\pi_{1i}}(N_{\pi_{1i}}) \geq \sum_{i=1}^{l_{\max}} w_i^{*\prime} \bar{X}_{\pi_{Vi}}(N_{\pi_{Vi}}).$$

Thus

$$\begin{aligned}
&\Pr\left\{ \max_{\mathbf{w}\in\mathscr{S}} \sum_{i=1}^{k} w_i' V_{\pi_{Vi}} > \max_{\mathbf{w}\in\mathscr{S}} \sum_{i=1}^{l_{\max}} w_i' \bar{X}_{\pi_{1i}}(N_{\pi_{1i}}) + z_{hi} B \right\} \\
&\leq \Pr\left\{ \sum_{i=1}^{l_{\max}} w_i^{*\prime} V_{\pi_{Vi}} > \sum_{i=1}^{l_{\max}} w_i^{*\prime} \bar{X}_{\pi_{Vi}}(N_{\pi_{Vi}}) + z_{hi} B \right\} \\
&= \Pr\left\{ \sum_{i=1}^{l_{\max}} w_i^{*\prime} (\bar{X}_{\pi_{Vi}}(N_{\pi_{Vi}}) - V_{\pi_{Vi}}) < -z_{hi} B \right\}.
\end{aligned}$$

The weighted average $A := \sum_{i=1}^{l_{\max}} w_i^{*\prime} (\bar{X}_{\pi_{Vi}}(N_{\pi_{Vi}}) - V_{\pi_{Vi}})$ is normal with mean zero and variance $\sum_{i=1}^{l_{\max}} (w_i^{*\prime})^2 \sigma^2_{\pi_{Vi}} / N_{\pi_{Vi}}$, where $\sigma_i^2 := \mathrm{Var}[X|Z_i]$. Furthermore,

$$\begin{aligned}
B^2 &= \max_{\mathbf{w}\in\mathscr{S}} \sum_{i=1}^{l_{\max}} (w_i')^2 S^2_{\pi_{Si}}(N_{\pi_{Si}}) / N_{\pi_{Si}} \\
&\geq \max_{\mathbf{w}\in\mathscr{S}} \sum_{i=1}^{l_{\max}} (w_i')^2 S^2_{\pi_{Vi}}(N_{\pi_{Vi}}) / N_{\pi_{Vi}} \\
&\geq \sum_{i=1}^{l_{\max}} (w_i^{*\prime})^2 S^2_{\pi_{Vi}}(N_{\pi_{Vi}}) / N_{\pi_{Vi}} := S^2.
\end{aligned}$$

The mean of S^2 is $\mathrm{E}[S^2] = \sum_{i=1}^{l_{\max}} (w_i^{*\prime})^2 \sigma^2_{\pi_{Vi}} / N_{\pi_{Vi}}$. The variance of S^2, based on the chi-squared distribution associated with sample variances, is $\mathscr{O}(N^{-3})$. For large N, this is negligible compared to the variance of A^2, which is $\mathscr{O}(N^{-2})$. Therefore, where Φ represents the standard normal cumulative distribution function,

$$\begin{aligned}
\Pr\{A < -z_{hi} B\} &= \frac{1}{2} \Pr\left\{A^2 > z_{hi}^2 B^2\right\} \\
&\leq \frac{1}{2} \Pr\left\{A^2 > z_{hi}^2 S^2\right\} \\
&\approx \frac{1}{2} \Pr\left\{A^2 > z_{hi}^2 \mathrm{E}[S^2]\right\} \\
&= \Pr\left\{A < -z_{hi} \sqrt{\mathrm{E}[S^2]}\right\} \\
&= \Phi\left(\frac{-z_{hi} \sqrt{\mathrm{E}[S^2]}}{\sqrt{\sum_{i=1}^{l_{\max}} (w_i^{*\prime})^2 \sigma^2_{\pi_{Vi}} / N_{\pi_{Vi}}}} \right) \\
&= \Phi(-z_{hi}) = \alpha_{hi}.
\end{aligned}$$

5 CONCLUSIONS

We have presented and tested a new two-level simulation procedure that creates a confidence interval for tail conditional expectation given a fixed computational budget. The outer level simulates risk factors and the inner level estimates portfolio values conditional on those risk factors. We are not aware of any research on confidence intervals generated by a simulation of this structure prior to Lan, Nelson, and Staum (2007). We found extremely low coverage for a confidence interval constructed by ignoring uncertainty from the inner level of simulation and using previous findings about the asymptotic variance of a point estimator for tail conditional expectation. We found that our procedure's coverage is greater than nominal when at least 1,000 risk factors were sampled, which enables the technique of empirical likelihood to provide adequate coverage. However, our procedure produces a rather wide confidence interval unless the computational budget C allows many millions of simulated payoffs, and we found that the width decreases very slowly in the budget, roughly as $\mathscr{O}(C^{-1/4})$.

Using more advanced efficiency techniques could decrease the width of the confidence interval. We tested a two-stage procedure without variance reduction techniques other than common random numbers in screening. Multistage screening, as used in Lesnevski, Nelson, and Staum (2006), should significantly improve efficiency. Variance reduction techniques applied to the inner level, such as the control variates used in Lesnevski, Nelson, and Staum (2006), will be helpful. Applying variance reduction techniques at the outer level might help, but it would be more difficult. Furthermore, some variance reduction techniques at the outer level would add little benefit given the existing

efficiency enhancements of our procedure. For example, importance sampling is often used in risk management simulations to increase the number of simulated risk factors that lead to large losses, but our procedure accomplishes something similar by screening out those risk factors that do not lead to large losses.

There may also be some ways to improve the speed of the procedure. The computational budget is expressed in terms of number of payoffs simulated, which should be the dominant computational cost in realistic examples, and ignores the time required to perform sorting, screening, and certain optimizations. Potential improvements include making screening faster and replacing the optimization that defines the factor B appearing in the upper confidence limit (2) with something simpler. When the inner-level sample sizes N_i are large, it might be possible to retain good coverage while replacing B with a quantity such as

$$\sqrt{\frac{1}{|I|}\sum_{i\in I}\frac{S_i^2(N_i)}{N_i}\max_{\mathbf{w}\in\mathscr{S}}\sum_{i=1}^{l_{\max}}(w_i')^2}$$

which would be easier to compute. Indeed, $\max_{\mathbf{w}\in\mathscr{S}}\sum_{i=1}^{l_{\max}}(w_i')^2$, which also features in the optimization of Equation (3), could be tabled instead of performing an optimization each time the simulation is run.

A greater potential improvement could come from tightening the lower confidence limit, which is currently based on controlling separately the difference of sample average and mean for all risk factors that survive screening. By contrast, the upper confidence limit is based on controlling a single weighted average of differences between sample averages and means. This would make the procedure less conservative, reducing the confidence interval width and the excessive coverage.

ACKNOWLEDGMENTS

This material is based upon work supported by the National Science Foundation under Grant No. DMI-0555485. We are grateful for the assistance of Jonathan Bragg, whose work was supported by a General Motors / EXCEL Alumni Undergraduate Research Grant from the Robert R. McCormick School of Engineering and Applied Science's Office of Corporate Relations.

REFERENCES

Baysal, R. E., and J. Staum. 2007. Empirical likelihood for value at risk and expected shortfall. Working Paper 07-01, Dept. of Industrial Engineering and Management Sciences, Northwestern University.

Bazaraa, M. S., H. D. Sherali, and C. M. Shetty. 2005. *Nonlinear Progarmming: Theory and Algorithms*, 3rd. ed. New York: John Wiley & Sons.

Björk, T. 1998. *Arbitrage Theory in Continuous Time*. New York: Oxford University Press.

Lan, H., B. L. Nelson, and J. Staum. 2007. Two-level simulations for risk management. In *Proceedings of the 2007 INFORMS Simulation Society Research Workshop*, ed. S. Chick, C.-H. Chen, S. Henderson, and E. Yücesan, 102–107.

Lee, S.-H. 1998. Monte Carlo computation of conditional expectation quantiles. Dissertation, Dept. of Operations Research, Stanford University.

Lesnevski, V., B. L. Nelson, and J. Staum. 2006. An adaptive procedure for estimating coherent risk measures based on generalized scenarios. Working Paper 06-05, Dept. of Industrial Engineering and Management Sciences, Northwestern University.

Manistre, B. J., and G. H. Hancock. 2005. Variance of the CTE estimator. *North American Actuarial Journal* 9: 129–154.

Owen, A. B. 2001. *Empirical Likelihood*. New York: Chapman & Hall/CRC.

AUTHOR BIOGRAPHIES

HAI LAN is a Ph. D. student in the Department of Industrial Engineering and Management Sciences at Northwestern University. His email address is <h-lan@northwestern.edu>.

BARRY L. NELSON is the Charles Deering McCormick Professor of Industrial Engineering and Management Sciences at Northwestern University. His research centers on the design and analysis of computer simulation experiments on models of stochastic systems, and he is Editor in Chief of *Naval Research Logistics*. Nelson has held many positions for the Winter Simulation Conference, including Program Chair in 1997 and current membership on the Board of Directors. His e-mail and web addresses are <nelsonb@northwestern.edu> and <www.iems.northwestern.edu/~nelsonb/>.

JEREMY STAUM is Associate Professor of Industrial Engineering and Management Sciences at Northwestern University. His research interests center on risk management and simulation in financial engineering. Staum is Associate Editor of *ACM Transactions on Modeling and Computer Simulation*, *Naval Research Logistics*, and *Operations Research*, and is Risk Analysis track coordinator at the 2007 Winter Simulation Conference. His e-mail address is <j-staum@northwestern.edu> and his web page is <www.iems.northwestern.edu/~staum>.