Proceedings of the 2006 Winter Simulation Conference
L. F. Perrone, F. P. Wieland, J. Liu, B. G. Lawson, D. M. Nicol, and R. M. Fujimoto, eds.

AN ADAPTIVE PROCEDURE FOR ESTIMATING COHERENT RISK MEASURES BASED ON GENERALIZED SCENARIOS

Vadim Lesnevski
Barry L. Nelson
Jeremy Staum

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

ABSTRACT

Coherent risk measures based on generalized scenarios can be viewed as estimating the maximum expected value from among a collection of simulated "systems." We present a procedure for generating a fixed-width confidence interval for this coherent risk measure. The procedure improves upon previous methods by being reliably efficient for simulation of generalized scenarios and portfolios with heterogeneous characteristics.

1 INTRODUCTION

Coherent risk measures can improve the practice of risk management (Artzner et al. 1999) and pricing derivative securities (Jaschke and Küchler 2001, Staum 2004). However, coherent risk measures may need to be estimated by simulation, which could be much slower than simulations currently used in risk management and derivatives pricing, too slow for routine use in practice.

Any coherent risk measure ρ has a representation of the form $\rho(Y) = \sup_{\mathbf{P}\in\mathcal{P}} \mathrm{E}_{\mathbf{P}}[-Y/r]$, where Y is the value of a portfolio at a future time horizon, r is a stochastic discount factor which represents the time value of money, and $\mathcal{P}$ is a set of probability measures (Artzner et al. 1999, Prop. 4.1). Equations of a similar form hold for the related problems in derivative security pricing. We simplify the problem somewhat by assuming that the set $\mathcal{P}$ has only a finite number k of elements $\mathbf{P}_1, \mathbf{P}_2, \ldots, \mathbf{P}_k$. This assumption often holds, for instance, when the decision maker designs the coherent risk measure (or the underlying acceptance set, in the case of derivative security pricing) by specifying k generalized scenarios. The assumption also covers approximation of $\mathcal{P}$ by the convex hull of k probability measures. Let $X := -Y/r$ and $\mu_i := \mathrm{E}_{\mathbf{P}_i}[X]$. The risk measurement involves a single random variable X, which is a negative discounted portfolio value or a discounted loss, viewed under multiple probability measures. For clarity in discussing simulations, let X_i be a random variable whose distribution under the probability measure $\Pr$ is the same as that of X under $\mathbf{P}_i$, that is, such that $\Pr[X_i \le x] = \mathbf{P}_i[X \le x]$. Because of the parallel with ranking and selection, we refer to X_i as an observation of system i.

In Lesnevski et al. (2005), we used tools from the ranking-and-selection literature to create efficient procedures for generating a fixed-width confidence interval for a coherent risk measure. These procedures use screening to eliminate some systems that seem likely to be inferior after generating a small number of observations. One of the procedures uses the technique of "restarting," in which the screening data are discarded so as to make it possible to reduce the required sample sizes for the systems that survive screening.

To sharpen screening we employ common random numbers (CRN; see Law and Kelton 2000) to induce positive correlation between the alternatives and thereby reduce the variance of their differences. To reduce the number of replications required for estimation, we employ control-variate estimators (CV; see Law and Kelton 2000) to exploit strong correlation between the response of interest, X, and a vector C of random variables with known expectations, called control variates.

A disadvantage of the procedures presented in Lesnevski et al. (2005) is that in some cases, they might require some previous knowledge about the problem to be efficient. For example, having a large screening budget is usually good, as it allows the procedure to screen out most of inferior systems. However, it might significantly decrease efficiency if more than one system has the maximum mean, or if some systems are nearly tied with the best. In such situations, screening might not be able to eliminate all systems but one. Even though the procedure with restarting is usually preferable over other alternatives, if screening is ineffective, restarting is wasteful of data. Before running the simulation, the user

1-4244-0501-7/06/$20.00 ©2006 IEEE

would have to decide whether or not to use restarting, and how much data to allocate to the screening stage. Making a good decision without substantial experience with simulation problems of the same form is difficult.

Without restarting, information generated during screening is reused during estimation of the confidence interval, so the amount of work done during screening is not very important. With restarting, information generated during screening is thrown away, so it is important to make sure that no excess work is done during screening. The advantage of restarting is that the new data are statistically independent of the screening exercise, so one may ignore the measures which were screened out, and design for the smaller problem.

In this paper we develop an adaptive multi-stage procedure which combines good features of both approaches. The procedure is very efficient for all configurations, as it gains the benefits of restarting and of having a large budget to use for screening.

2 ADAPTIVE MULTI-STAGE PROCEDURE

Our procedure produces a lower confidence limit that covers the coherent risk measure with probability at least $1-\alpha_a$, and an upper confidence limit that covers with probability at least $1-\alpha_b$. See Lesnevski (2006) for a proof. The procedure spends some of this allowable error on screening (α_I), some on control variates (α_C), and the remainder on estimating the means of some systems. We use the control variate C_i for the output X_i of system i to improve estimation of the mean μ_i of system i.

The adaptive multi-stage procedure consists of two phases. Phase I ("pre-screening") consists of multi-stage screening whose purpose is, while controlling relative cost, to screen out as many inferior systems as possible, so that they do not contribute to the critical values that determine the overall sample size for mean estimation. No samples obtained during pre-screening are used during Phase II, which is an estimation procedure with additional multi-stage screening.

2.1 Phase I: Pre-screening

The sole purpose of the first phase is to reduce the number of systems and thus the natural bias of the estimation problem, making a fixed-width confidence interval attainable with fewer replications.

The maximal number of Phase I stages, m, is specified in advance. The first stage of Phase I is stage 0 and the first stage of Phase II is stage M, where the random variable $M \leq m$. The decision to proceed to Phase II is made randomly, on the basis of the simulated data, when the cost of continuing and doing one more stage of Phase I is greater than the estimated approximate savings due to further pre-screening. The growth rate R and the initial sample size n_0 are also specified in advance, so that the total sample size during stage ℓ is $N(\ell) = \lceil n_0 R^\ell \rceil$.

The initial sample size n_0 should be chosen so that sample averages are approximately normal. In most cases, $n_0 = 30$ is adequate. The procedure is most efficient if the growth factor R is between 1.2 and 2.0, while m is such that the total budget available for pre-screening is large. For example, if $R = 1.5$ and $m = 30$, the total budget available for Phase I is $\lceil n_0 R^{m-1} \rceil = 3,835,021$, which is large enough for most applications. We found that $R = 1.5$ and $m = 30$ worked well on all problems we consider. It wasn't possible to improve on the performance much by altering the parameters, as it was for the procedures presented in Lesnevski et al. (2005).

Let I be the set of systems that have not been screened out. Initially set $I \leftarrow \{1,\ldots,k\}$. Each stage $\ell = 1, \ldots, m$ of Phase I consists of the following steps:

1. *Simulation.*
   Simulate (X_{ij}, C_{ij}) for $j = N(\ell-1)+1, \ldots, N(\ell)$ and all $i \in I$.
2. *Screening.*
   For each $h, i \in I$ such that $h \neq i$, set

$$\begin{aligned}
\bar{\bar{D}}_{hi} &\leftarrow \frac{1}{N(\ell)} \sum_{j=1}^{N(\ell)} (X_{hj} - X_{ij}), \\
S^2_{hi} &\leftarrow \frac{1}{N(\ell)-1} \sum_{j=1}^{N(\ell)} (X_{hj} - X_{ij} - \bar{\bar{D}}_{hi})^2, \\
W_{hi} &\leftarrow \frac{t_{N(\ell)-1,1-\alpha_I/(2m(k-1))}}{\sqrt{N(\ell)}} S_{hi},
\end{aligned}$$

   where $t_{\nu,p}$ is the p quantile of the t distribution with ν degrees of freedom.
   Then set $I \leftarrow \left\{ i \in I | \forall h \in I, \bar{\bar{D}}_{hi} \geq -W_{hi} \right\}$.
3. *Checking whether to proceed to Phase II.*
   For each $i \in I$, compute the residual variance $\hat{\sigma}_i^2$ of regressing $X_{i,1}, \ldots, X_{i,N(\ell)}$ on $C_{i,1}, \ldots, C_{i,N(\ell)}$ and define

$$\begin{aligned}
c_p &:= \frac{1}{L}(\Phi^{-1}(1-\alpha_a/p+\alpha_C) + \\
&\qquad \Phi^{-1}(1-\alpha_b+\alpha_I+\alpha_C)), \qquad (1)
\end{aligned}$$

   where Φ is the standard normal cumulative distribution function. If

$$|I| N(\ell)(R-1) > (c^2_{|I|} - c^2_1) \max_{i \in I} \hat{\sigma}_i^2, \qquad (2)$$

   the procedure jumps to Phase II by setting $M \leftarrow \ell+1$, which means that the next stage is the first

stage of Phase II, and by setting $K \leftarrow |I|$, which is the number of systems left after pre-screening and which will be used for determining final sample sizes. Otherwise, set $\ell \leftarrow \ell+1$ and return to Step 1.

Under the transition rule given by Inequality (2), pre-screening stops after stage $M-1$ when the cost of doing one more stage of pre-screening is greater than the approximate maximal savings due to continuation, computed under the assumption that after additional pre-screening there will be only one system left and it will have the largest variance.

2.2 Phase II: Screening and Estimation

Phase II begins by restarting, that is, throwing out all the data obtained in Phase I. The only effect of Phase I on Phase II is that Phase I determines the subset I of systems that Phase II handles. Phase II contains three parts.

First, in the initial stage M, the procedure determines the required total sample sizes N_i for each of the systems in I and the maximal necessary number P of subsequent screening stages. Second, in stages $M, \ldots, M+P-1$, the procedure does more screening. It maintains two sets of systems: the set I contains systems that have survived screening and from which the procedure has simulated as many samples as are required to construct the fixed-width confidence interval, while the set $\hat{I}$ contains systems that have survived screening so far, but which still require more sampling. Finally, once the required sample size has been reached for all surviving systems, the procedure constructs a confidence interval.

Because M is the first stage after restarting, the procedure discards $\lceil n_0 R^{M-1} \rceil$ Phase I samples. To compensate for the discarded samples and keep the growth rate constant, during Phase II the procedure sets $N(\ell) \leftarrow n_0 R^{\ell-1}(R+1)$, $\ell \geq M$. This makes the total Phase II sample size grow at the rate R. It also makes the initial sample size of Phase II be $N(M)-N(M-1) \simeq n_0 R^M$, which is large enough to ensure high-quality variance estimates.

Initialize $\hat{I} \leftarrow I$ and then $I \leftarrow \emptyset$. Also initialize $N_i \leftarrow N(M)$ for all $i \in \hat{I}$. Each stage $\ell = M, \ldots, M+P$ consists of the following steps, except that only stage M contains Step 2, and Step 4 will not occur during stage $M+P$ because $\hat{I}$ will be empty then:

1. *Simulation.*
   Simulate (X_{ij}, C_{ij}) for $j = N(\ell-1)+1, \ldots, \min\{N_i, N(\ell)\}$ and all $i \in \hat{I}$.
   Set $n \leftarrow N(\ell)-N(M-1)$.
2. *Setting final sample sizes.*
   If $\ell > M$, skip this step.
   Set $\alpha_a'' \leftarrow \alpha_a/K - \alpha_C$ and $\alpha_b'' \leftarrow \alpha_b - \alpha_I - \alpha_C$, and set the scaling constant

$$c \leftarrow \frac{1}{L}(t_{n-q-1,1-\alpha_a''} + t_{n-q-1,1-\alpha_b''}), \qquad (3)$$

   where $q := \max_{i \in I} q_i$ and each q_i is the number of control variates in C_i.
   For each $i \in \hat{I}$, compute the residual variance $\hat{\sigma}_i^2$ of regressing $X_{i,N(M-1)+1}, \ldots, X_{i,N(M)}$ on $C_{i,N(M-1)+1}, \ldots, C_{i,N(M)}$, and from it the total sample size

$$N_i \leftarrow \lceil c^2 \hat{\sigma}_i^2 + \chi^2_{q_i,1-\alpha_C} \rceil + N(M-1), \qquad (4)$$

   where $\chi^2_{\nu,p}$ is the p quantile of the chi-squared distribution with ν degrees of freedom.
   Set $P \leftarrow \lceil \log_R \max_{i \in I}(N_i/N(M)) \rceil$.
3. *Updating I and $\hat{I}$.*
   Add to I systems that have reached their required sample sizes and remove them from $\hat{I}$: set $I \leftarrow I \bigcup \{i \in \hat{I} | N_i \leq N(\ell)\}$ and $\hat{I} \leftarrow \hat{I} \backslash I$.
4. *Screening.*
   For each $h, i \in \hat{I}$ such that $h \neq i$, set

$$\begin{aligned}
\bar{\bar{D}}_{hi} &\leftarrow \sum_{j=N(M-1)+1}^{N(\ell)} \frac{X_{hj}-X_{ij}}{n}, \\
S_{hi}^2 &\leftarrow \sum_{j=N(M-1)+1}^{N(\ell)} \frac{(X_{hj}-X_{ij}-\bar{\bar{D}}_{hi})^2}{n-1}, \\
W_{hi} &\leftarrow \frac{1}{\sqrt{n}} t_{n-1,1-\alpha_I/(2P(K-1))} S_{hi}.
\end{aligned}$$

   Then set $\hat{I} \leftarrow \left\{ i \in \hat{I} | \forall h \in I, \bar{\bar{D}}_{hi} \geq -W_{hi} \right\}$.
5. *Continue or compute confidence interval.*
   If $\hat{I} \neq \emptyset$, set $\ell \leftarrow \ell+1$ and return to Step 1.
   Otherwise, for each $i \in I$, compute the estimate $\widehat{\mu}_i$ from the regression of $X_{i,N(M-1)+1}, \ldots, X_{i,N_i}$ on $C_{i,N(M-1)+1}, \ldots, C_{i,N_i}$. Set

$$\begin{aligned}
a &\leftarrow \frac{1}{c} t_{N(M)-N(M-1)-q-1,1-\alpha_a''} \quad \text{and} \\
b &\leftarrow \frac{1}{c} t_{N(M)-N(M-1)-q-1,1-\alpha_b''}.
\end{aligned}$$

   The confidence interval is

$$(\max_{i \in I} \widehat{\mu}_i - a, \max_{i \in I} \widehat{\mu}_i + b).$$

2.3 Efficiency of the Rule for Restarting

The adaptive procedure offers two significant improvements over our previous procedures.

First, we do not need to specify a screening budget in advance. Choosing the screening budget too small or too big could have a very significant effect on the performance of our previous procedures, in some configurations making a simulation dozens of times slower: see Table 3 in Section 3. The adaptive procedure solves this problem by trying to screen out a system in Phase II only until its required sample size is reached. In effect, this allows the screening budget to be arbitrarily large, to vary by system, and to be determined adaptively by the required sample size.

Second, the adaptive procedure allows us to restart whatever the configuration of the means $\mu_1, \ldots, \mu_k$ may be. The effect of the decision whether or not to restart on performance is much less severe: as we will show below, usually we do not expect to save more than 40-80%. Restarting is usually beneficial because in a typical case there is only one best system. Having an adaptive pre-screening phase identifying a good time to restart allows us to achieve very good performance in a typical case, and reasonably good performance in all other cases.

How big are the benefits of pre-screening in a typical case? To answer this question let us first estimate the maximal possible savings due to restarting.

In the following analysis we make several simplifying assumptions. First, we assume that the estimate of the residual variance $\hat{\sigma}_i^2$ of system i is always approximately equal to the true residual variance σ_i^2. Second, we ignore the effect of the number of degrees of freedom on the sample sizes for estimation. Third, we assume that the effort required for screening out an inferior system is always the same, whether in Phase I, Phase II, or in an alternative procedure without pre-screening and restarting (such as the multi-stage procedure with early stopping of Lesnevski et al. 2005).

The total cost E of a simulation without pre-screening is the sum of the cost E_s of screening out inferior systems and the cost E_e of estimation of the surviving systems: $E = E_s + E_e$.

The total cost $\tilde{E}$ of a simulation with pre-screening is the sum of the pre-screening cost $\tilde{E}_p$, the cost $\tilde{E}_s$ of screening out inferior systems in Phase II, and the estimation cost $\tilde{E}_e$ of the surviving systems: $\tilde{E} = \tilde{E}_p + \tilde{E}_s + \tilde{E}_e$.

Under our assumptions, the sample size N_i in Equation (4) is approximately equal to $c^2\sigma_i^2$. Without pre-screening, the constant c in Equation (3) is approximately equal to c_k defined in Equation (1), where k is the initial number of systems. With pre-screening, c is approximately c_K, where K is the number of systems remaining after pre-screening. The smaller K, the bigger the benefit of pre-screening, because smaller c_K leads to smaller sample sizes for estimation.

We will assume that whether we simulate with pre-screening or not, the set I of the surviving systems is the same. This is generally so when pre-screening is stopped before the sample sizes for some systems exceed the sample sizes required for estimation, which is exactly the case when pre-screening could be beneficial.

A simulation without pre-screening costs $E = E_s + c_k^2 \sum_{i \in I} \sigma_i^2$, and a simulation with pre-screening costs $\tilde{E} = \tilde{E}_p + \tilde{E}_s + c_K^2 \sum_{i \in I} \sigma_i^2$. The latter is minimized when c_K^2 is as small as possible, which occurs when $K = 1$, i.e., there is only one system left after pre-screening. Also, under the assumptions we use in this section, the screening cost E_s is less than the total of the pre-screening and screening costs $\tilde{E}_p + \tilde{E}_s$, so the maximal efficiency improvement $E/\tilde{E}$ is achieved when the pre-screening and screening costs are negligible compared to estimation costs. This is a typical case in practice: pre-screening and screening are very fast compared to estimation, and they eliminate all but one system. Under our assumptions, and if pre-screening and screening costs are negligible, the efficiency improvement due to restarting (i.e., due to having a pre-screening phase) is

$$\frac{E}{\tilde{E}} \approx \frac{c_k^2 \sum_{i \in I} \sigma_i^2}{c_K^2 \sum_{i \in I} \sigma_i^2} = \frac{c_k^2}{c_K^2} \leq \frac{c_k^2}{c_1^2}.$$

Figure 1 shows the maximal efficiency improvement c_k^2/c_1^2 as a function of the initial number of systems k. When the number of systems k is between 20 and 1000, the savings in a typical case are 40-80% at $1-\alpha = 99\%$ confidence and 60-140% at $1-\alpha = 95\%$ confidence.

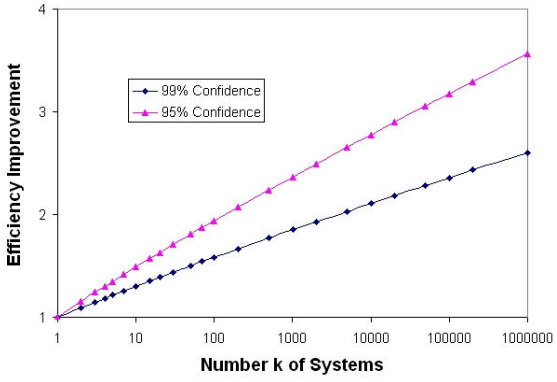

Figure 1: Maximal Efficiency Improvement Due to Restarting with $\alpha_a = 0.8\alpha$ and $\alpha_b = 0.2\alpha$

Recall that the transition rule given by Inequality (2) chooses to restart when the cost of doing one more stage of pre-screening is greater than the approximate maximal savings due to continuation, computed under the assumption that after additional pre-screening there will be only one system left and it will have the largest variance. A typical case indeed has one clear best system, so the effort required for screening out inferior systems is relatively small, the

approximate maximal savings are relatively large, and pre-screening makes I a singleton.

How efficient is this transition rule in other situations? Let us consider a configuration when there are several systems which are tied for the best, while other systems are relatively easy to screen out. In this case we might worry that the cost of pre-screening could get too high before the adaptive procedure proceeds to Phase II. Is our transition rule still efficient?

Because now we are concerned that pre-screening may be too expensive, we assume that pre-screening lasts a long time and eliminates all inferior systems: the set $I(M)$ of systems used in Phase II equals I, the set of systems that survive screening and reach their required sample sizes, and the Phase II cost of screening $\tilde{E}_s = 0$. Again we assume that I is the same whether we use pre-screening or not: here we assume it contains only the systems that are tied. We now show how the transition rule in Inequality (2) provides a bound on $\tilde{E}_p - E_s$, the excess cost of pre-screening in the adaptive procedure over the cost of screening in a procedure without restarting. The effort required to screen out inferior systems is similar in either procedure, so $\tilde{E}_p - E_s \approx KN(M-1)$, the number of samples from the $K = |I|$ surviving systems that the adaptive procedure throws out by restarting.

Pre-screening stops after stage $\ell = M-1$, the first time that the cost $(R-1)|I(\ell+1)|N(\ell)$ of the next stage exceeds $(c^2_{|I(\ell+1)|} - c_1^2)\max_{i\in I(\ell+1)} \hat{\sigma}_i^2(\ell)$. Under our present assumption that the residual variance estimates are approximately correct, this yields the approximate upper bound

$$\begin{aligned} N(M-2) &\leq \frac{(c^2_{|I(M-1)|} - c_1^2)\max_{i\in I(M-1)} \sigma_i^2}{(R-1)|I(M-1)|} \\ &\leq \frac{(c_K^2 - c_1^2)\max_{i\in I} \sigma_i^2}{(R-1)K} \end{aligned}$$

because $I(M-1)$ contains $I(M) = I$ whose size is K, and c_p^2 defined in Equation (1) increases in p at a rate that is less than linear. Thus

$$\begin{aligned} KN(M-1) &\leq KRN(M-2) \\ &\leq \frac{R(c_K^2 - c_1^2)\max_{i\in I} \sigma_i^2}{R-1}. \end{aligned}$$

For $R = 1.5$, $R/(R-1) = 3$, and the relative efficiency improvement is

$$\begin{aligned} \frac{E}{\tilde{E}} &= \frac{E_s + c_k^2 \sum_{i\in I} \sigma_i^2}{\tilde{E}_p + \tilde{E}_s + c_K^2 \sum_{i\in I} \sigma_i^2} \\ &= \frac{E_s + c_k^2 \sum_{i\in I} \sigma_i^2}{E_s + 3(c_K^2 - c_1^2)\max_{i\in I} \sigma_i^2 + c_K^2 \sum_{i\in I} \sigma_i^2} \\ &\approx \frac{c_k^2 \sum_{i\in I} \sigma_i^2}{3(c_K^2 - c_1^2)\max_{i\in I} \sigma_i^2 + c_K^2 \sum_{i\in I} \sigma_i^2} \end{aligned}$$

approximately, if the cost E_s of screening is small. If the variances of the tied systems are approximately equal, this simplifies to

$$\frac{Kc_k^2}{3(c_K^2 - c_1^2) + Kc_K^2}.$$

For $k = 256$ and $k = 64$ the efficiency improvements as a function of the number K of tied systems are shown in Figure 2. A value less than 1 represents a loss of efficiency. We see that even when some systems are tied, restarting with our transition rule can still produce substantial benefits. Even when all the systems are tied, the loss of efficiency is very slight.

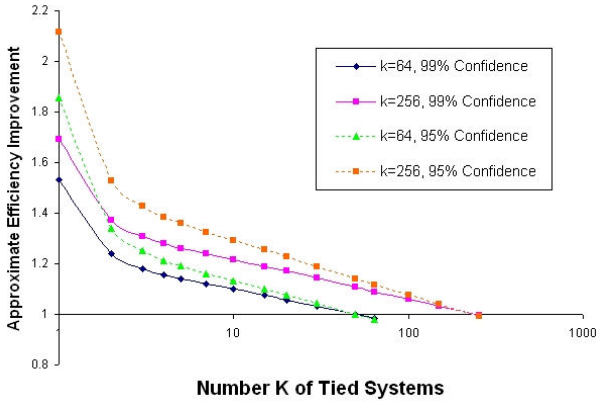

Figure 2: Effect of Ties on Approximate Efficiency Improvement Due to Restarting with $\alpha_a = 0.8\alpha$ and $\alpha_b = 0.2\alpha$

The transition rule we presented is heuristic and is one of many similar rules that all work well. This rule is advantageous because of its simplicity and because it allows us to reap most of the benefits of restarting, without causing significant inefficiencies when restarting could be harmful. More efficient transition rules could be designed which take into account not only the sample variances of the systems, but also their sample means. However, such rules are complicated, and in most cases provide either small or no savings. Because the benefits seem insufficient to justify the additional complexity, we do not consider this approach here.

3 PERFORMANCE OF THE ADAPTIVE MULTI-STAGE PROCEDURE

In this section we use the basket put and options portfolio examples discussed in Lesnevski et al. (2005) to illustrate our procedure.

The basket put is the right to sell a certain number of shares of three securities for a fixed price at a future date. The simulation problem is to find the maximum

expectation of the resulting discounted payoff, where the expectation can be taken under any of $k = 64$ probability measures. The 64 probability measures differ from each other in the correlations among the three securities. The basket put example uses put options on each of the individual securities in the basket as control variates.

The options portfolio example measures the risk of a portfolio of put and call options on three securities. The simulation problem is to find the maximum conditional expectation of the portfolio's loss, where the conditional expectation can be taken under any of $k = 256$ generalized scenarios. The generalized scenarios are constructed by conditioning on events in which some of four underlying risk factors attain high, low, or moderate values. The options portfolio example does not use control variates.

To test the adaptiveness of the procedure, in addition to the ordinary configuration with one best system, we also consider configurations "2 best" (obtained by adding a duplicate of the best system), "4 best" (by adding 3 duplicates), and "16 best" (by adding 15 duplicates), so that configuration "2 best" in the basket put example has $64+1 = 65$ systems in total, while configuration "16 best" has $64+15 = 79$ systems. This is not the same as in Figure 2, where the total number k of systems remains constant while the number K that are tied varies.

We split the $1-\alpha = 1\%$ allowable error into components $\alpha_a = 0.8\%$ for the lower confidence limit and $\alpha_b = 0.2\%$ for the upper confidence limit. The error allocated to screening is $\alpha_I = 0.04\%$, and when using control variates, $\alpha_C = 0.002\%$ is allocated to controlling them. We choose initial sample size n_0 and the maximal number m of Phase I stages to be 30, and the growth factor R to be 1.5. We use CRN in all examples.

For ease of interpretation, we specify the fixed confidence interval width L as a percentage of a quantity which provides a natural scale for the example. For the options portfolio example, this quantity is the portfolio's standard deviation. For the basket put example, this quantity is the true value, the largest mean.

We report efficiency as a speed improvement relative to the standard procedure, a modification of the two-stage procedure of Chen and Dudewicz (1976), as explained in Lesnevski et al. (2005, §§4.1–4.2). That is, we report the ratio of the average number of samples required by the standard procedure to the average number of samples required by the adaptive multi-stage procedure. The results are summarized in Table 1. Recall that efficiency improvement can be larger than the number of systems k, which is 64 for the ordinary configuration of the basket put, and 256 for that of the options portfolio. The reason is that the improvement depends not only on k, but also on the size of the best system's standard deviation relative to the standard deviations of other systems.

Table 2 shows how much work the procedure does in excess of the work required by the "clairvoyant" procedure,

Table 1: Efficiency Relative to the Standard Procedure at 99% Confidence

Config.	Example					
	Options Portfolio			Basket Put		
	Precision					
	0.3%	1%	5%	0.3%	1%	5%
1 best	**252**	**244**	**154**	**208**	**158**	**22**
2 best	104	98	81	85	76	19
4 best	51	48	43	40	38	15
16 best	12	12	12	11	10	6.7

Table 2: Sample Size Relative to the Clairvoyant Procedure at 99% Confidence

Config.	Example					
	Options Portfolio			Basket Put		
	Precision					
	0.3%	1%	5%	0.3%	1%	5%
1 best	**1.0**	**1.1**	**1.7**	**1.1**	**1.4**	**10**
2 best	1.2	1.2	1.5	1.2	1.3	5.4
4 best	1.1	1.2	1.3	1.2	1.2	3.2
16 best	1.1	1.1	1.1	1.1	1.1	1.7

the procedure that knows in advance which systems are tied for the best, and applies the standard procedure to only these systems in isolation. That is, the clairvoyant procedure screens out all inferior systems by guessing right with no work.

Like the multi-stage procedure with restarting analyzed in Lesnevski et al. (2005), the adaptive procedure is less than 10% more expensive than estimating a single mean in the "1 best" configuration when a precise estimate is required. If there are ties the procedure first tries to break them, but when this becomes too expensive, proceeds to estimation: this is its advantage over the multi-stage procedure with restarting. Table 2 demonstrates the robustness of the adaptive procedure's performance to configuration.

As we see from the last column of Table 2, in the configuration with no ties at 5% precision the adaptive procedure looks relatively inefficient compared to the clairvoyant procedure (10 times slower), but adding ties can make the adaptive procedure look more favorable. This is because 5% is a low precision, so the final sample size is not very large relative to the sample size required for screening. At 5% precision the clairvoyant procedure has a big advantage in screening perfectly for free.

Table 3 shows the efficiency improvement of the adaptive procedure relative to the most efficient procedure of Lesnevski et al. (2005): the multi-stage procedure with restarting. (In all cases reported in Table 3, the multi-

Table 3: Efficiency Relative to the Multi-Stage Procedure with Restarting at 99% Confidence

Configuration and Precision		Options Portfolio: Number of screening stages m = 5	10	15	20	25	30	Basket Put: Number of screening stages m = 5	10	15	20	25	30
1 best	0.3%	1.0	1.0	1.0	1.0	1.0	1.0	41	2.8	1.0	1.0	1.0	1.0
	1%	1.0	1.0	1.0	1.0	1.0	1.0	30	2.3	1.0	1.0	1.0	1.0
	5%	1.0	1.0	1.0	1.0	1.0	1.0	4.7	1.0	0.9	0.9	1.0	1.0
2 best	0.3%	0.9	0.9	0.9	0.9	1.0	1.8	17	1.6	0.9	1.0	1.6	6.1
	1%	0.9	0.9	0.9	1.0	2.0	10	16	1.5	1.0	1.8	7.8	54
	5%	0.9	0.9	1.2	4.0	25	205	4.6	1.0	1.6	6.5	44	328
4 best	0.3%	0.9	0.9	0.9	0.9	1.0	1.8	8.3	1.3	0.9	1.0	1.5	5.8
	1%	0.9	0.9	0.9	1.0	2.1	10	7.8	1.2	1.0	1.8	7.7	53
	5%	0.9	0.9	1.2	4.1	27	197	3.4	1.1	2.0	10	68	509
16 best	0.3%	0.9	0.9	0.9	0.9	1.0	1.7	3.1	1.0	1.0	1.0	1.6	5.7
	1%	0.9	0.9	1.0	1.1	2.0	9.2	3.0	1.1	1.1	1.8	7.7	52
	5%	0.9	1.0	1.2	4.3	27	201	2.1	1.1	2.7	15	110	833

stage procedure with early stopping was somewhat more expensive than the multi-stage procedure with restarting.) In some cases, the efficiency is slightly less than 1, i.e., the adaptive procedure required slightly more samples than the multi-stage procedure with restarting: the adaptive procedure does not always pick the best possible time to restart, but it picks a good time.

The efficiency of both of the procedures depends heavily on the actual configuration of the means and the total screening budget of $n_0 R^{m-1}$ observations per system. We tested these procedures with $n_0 = 30$ and $R = 1.5$ while varying the maximal number of stages available for screening from 5 to 30, so that the total budget available for screening varied from 152 to $3,835,022$ observations per system. We set $R = 1.5$, not $R = 2$ as in Lesnevski et al. (2005), as this choice of the growth factor makes all procedures more efficient when there are ties.

The results in Table 3 illustrate the danger for our previous multistage procedures of choosing the budget for screening either too small or too large. What constitutes too small or too large depends on the actual configuration, whereas the adaptive procedure works well in all of them.

ACKNOWLEDGMENTS

This material is based upon work supported by the National Science Foundation under Grants No. DMI-0217690, DMS-0202958, and DMI-0555485.

REFERENCES

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

Chen, H. J., and E. J. Dudewicz. 1976. Procedures for fixed-width interval estimation of the largest normal mean. *Journal of the American Statistical Association* 71: 752–756.

Jaschke, S., and U. Küchler. 2001. Coherent risk measures and good-deal bounds. *Finance and Stochastics* 5: 181–200.

Law, A. M., and W. D. Kelton. 2000. *Simulation Modeling and Analysis*, 3rd ed. McGraw-Hill, New York.

Lesnevski, V. 2006. Simulation of coherent risk measures based on generalized scenarios. Ph. D. dissertation, Dept. of Industrial Engineering and Management Sciences, Northwestern University.

Lesnevski, V., B. L. Nelson, and J. Staum. 2005. Simulation of coherent risk measures based on generalized scenarios. Working paper 05-002, Dept. of Industrial Engineering and Management Sciences, Northwestern University.

Nelson, B. L., and J. Staum. 2006. Control variates for screening, selection, and estimation of the best. *ACM Transactions on Modeling and Computer Simulation* 16: 52–75.

Staum, J. 2004. Fundamental theorems of asset pricing for good deal bounds. *Mathematical Finance* 14: 141–161.

AUTHOR BIOGRAPHIES

VADIM LESNEVSKI is a Ph. D. student in the Department of Industrial Engineering and Management Sciences at Northwestern University. His email address is <v-lesnevski@northwestern.edu>.

BARRY L. NELSON is the Charles Deering McCormick Professor of Industrial Engineering and Management Sci-

ences at Northwestern University, and is Director of the Master of Engineering Management Program there. 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 Assistant Professor of Industrial Engineering and Management Sciences at Northwestern University. His research interests include variance reduction, financial simulations, and pricing derivative securities in incomplete markets. Staum is Associate Editor of *ACM Transactions on Modeling and Computer Simulation* and *Naval Research Logistics*. His e-mail address is <j-staum@northwestern.edu> and his web page is <www.iems.northwestern.edu/~staum>.