A Fast Monte Carlo Algorithm for Estimating Value at Risk and Expected Shortfall

MING-HUA HSIEH, WEI-CHENG LIAO, AND CHUEN-LUNG CHEN

MING-HUA HSIEH
is an associate professor in the Department of Risk Management and Insurance at National Chengchi University in Taiwan.
mhsieh@nccu.edu.tw

WEI-CHENG LIAO
is a Ph.D candidate in the Department of Management Information Systems at National Chengchi University in Taiwan.
97356505@nccu.edu.tw

CHUEN-LUNG CHEN
is a professor in the Department of Management Information Systems at National Chengchi University in Taiwan.
chencl@nccu.edu.tw

A reliable estimation of loss probabilities is essential for calculating value at risk and expected shortfall. Typically, a full valuation approach is favored in the estimation process because it yields a very accurate price of financial instruments. Monte Carlo simulation is often used in conjunction with a full valuation approach because the market scenarios can be generated from the stochastic model of the risk factors; however, the computational cost of a full valuation approach is high. In this article, a fast importance sampling procedure is proposed for estimating loss probabilities using a full valuation approach. The foundation of the algorithm uses a special matrix decomposition technique, the spectral decomposition, which identifies the key factor that drives the dependence structure of the risk factors. Furthermore, the monotonic property between the key factor and the portfolio loss exceeding the loss threshold provides the guidelines for selection of appropriate importance sampling distributions. In particular, the technique of zero-variance changes of measure is adopted for selection of importance sampling distribution for the key factor. The likelihood ratio of the importance sampling estimator is less than one for every sample path generated, and therefore, we guarantee the algorithm facilitates variance reduction over naive simulation. The performance improves immensely when moving the loss threshold to more extreme values or when taking the dependence structure to more correlated situations. Numerical results indicate that the estimator exhibits a constant coefficient of variation that suggests the importance sampling estimator has bounded relative error.

Value at risk (VaR) and expected shortfall (ES) are both popular risk management tools applied to market exposure to ensure that banks have sufficient capital during stressful market events. VaR describes the threshold of portfolio losses for a given probability as a single value for measuring risk (Wilson [1999], Jorion [2007], Alexander [2009], Hull [2012]). Artzner et al. [1997, 1999] pointed out that any risk measure that satisfies the four properties—drift invariance, homogeneity, monotonicity, and subadditivity—is known as coherent. However, VaR is not a coherent risk measure because it fails to fulfill the subadditivity criterion. ES is a widely used alternative coherent risk measure that evaluates portfolio average losses exceeding VaR.

Comprehensively, a reliable risk measure relies on accurate estimation of portfolio loss, and the calculation is relatively straightforward if given the loss distribution. However, estimating a loss distribution can be time-consuming because portfolio losses are aggregated from the return of each instrument. Generally, the behaviors of risk factors are based on specific market models. Monte Carlo simulation is the predominant numerical method used to define different market models because it exhibits parametric flexibility to assimilate to different market conditions. In Monte Carlo simulation, risk factor scenarios are generated from more elaborate market models, for example, a Brownian motion calibrated to historical covariance of instrument returns. However, Monte Carlo simulation is also the most computationally intensive method, in particular, for calculation of VaR and ES, because a large loss is a rare event. This article proposes a fast Monte Carlo simulation algorithm that improves the efficiency of estimating VaR and ES.

Let $V(t)$ be the value of the portfolio at time t, then define $V(t)$ as a function of the risk factors $S(t)$, thus $V(t) = V(S(t))$. Set the holding period to be Δt and the value of the portfolio at time $t + \Delta t$ be $V(t + \Delta t)$. This article defines the loss in the portfolio value during the holding period as follows:

$$L = V(t) - V(t+\Delta t) = V(S(t)) - V(S(t+\Delta t)) \qquad (1)$$

VaR is used to measure the magnitude of the loss at a specific confidence level $(1 - \alpha)$, such as 95% or 99%. For a given probability α, VaR x_α is defined as

$$P(L > x_\alpha) = \alpha \qquad (2)$$

x_α is the $(1 - \alpha)$th percentile of the loss distribution; thus, the estimation of x_α presents a quantile estimation problem. A direct simulation method for generating quantile estimations must consider the loss distribution. However, it is convenient to compute the large loss probability and subsequently estimate the quantile (Glasserman et al. [2000]). In this article, the problem of interest is to estimate the large loss probability p, when considering a large loss threshold b. Define

$$p = P(L > b) \qquad (3)$$

It is obvious that if an efficient method is used to compute p for any b, then the calculation of x_α is straightforward.

From a statistical perspective, x_α is defined as the $(1 - \alpha)$ th quantile associated with the loss distribution L, and ES is the conditional expectation of the truncated loss distribution at the confidence level α. Define

$$E[L \mid L > x_\alpha] = \frac{E\left[L; L > x_\alpha\right]}{P(L > x_\alpha)} \qquad (4)$$

where the denominator is the large loss probability. To calculate the ES, we must determine the value of the numerator, namely, expected loss above threshold, $E[L;\ L > x_\alpha]$. Because x_α is unknown, it is difficult to generate a direct simulation method to calculate $E[L;\ L > x_\alpha]$. An accurate estimation of $E[L;\ L > x_\alpha]$ requires numerous replications for the tail distribution. As a result, rather than estimating precise x_α and loss distribution, $E[L; L > b]/P(L > b))$ yields a good approximation of ES when $P(L > b) \approx \alpha$. Thus, this article focuses on developing a fast algorithm for evaluating $E[L; L > b]$ and $P(L > b)$ for any given loss threshold b. In particular, without generating a complete loss distribution, this article contributes a direct simulation algorithm for accurately estimating $E[L; L > b]$ and $P(L > b)$ by Monte Carlo simulation.

The central idea of VaR estimation is to model the dynamic behavior of the risk factors to which the portfolio is exposed and to translate the changes in the risk factors into the changes in portfolio values. The classic VaR estimation method, the RiskMetrics method proposed by J.P. Morgan (Longerstaey [1996]), assumes the change in the value of the portfolio is driven by correlated risk factors, and thus portfolio value is a linear combination of the risk factors. The RiskMetrics method also assumes that the change in the value of the portfolio is normally distributed so that the VaR can be easily estimated, in which the required parameter is the covariance matrix of the risk factors. However, two factors define the assumption's shortcomings: First, portfolio value is not always a perfectly linear function of the risk factors, for example, in the case of option portfolios. Second, the risk factors exhibit a heavy tailed distribution rather than a normal distribution in real markets.

To generalize the model setting in the RiskMetrics method, literature on VaR estimations can be separated into two streams: One stream targets relaxing of model assumptions on the portfolio value, which is not a linear combination of the risk factors. In this stream, a specific quadratic function is widely used for approximation of portfolios losses. The approximation method is known as the delta-gamma approach based on the Taylor series expansion method to include quadratic effects (Rouvinez [1997]; Wilson [1999]). Yueh and Wong [2010] applied the Fourier transformation technique to derive an analytical expression for VaR and ES dependent on quadratic approximation of portfolio loss. Based on historical observations, however, another literature stream has considered that the true market is more volatile than is theoretically expected and the market returns exhibit a skewed, leptokurtic, or heavy-tailed distribution. Hull and White [1998] relaxed the normal distribution assumption by proposing a transformation model for VaR estimation, in which the risk factors can be any distribution, but the model requires the transformed new risk factors to follow a multivariate normal distribution.

This article contributes to both streams of literature. Instead of using an approximation to solve the nonlinear relationship between portfolio value and the risk factors, the full valuation evaluates accurate portfolio value at each time point. Typically, the full valuation is an optimal choice for calculation of VaR and ES because it yields a very accurate price for the instruments. Through accurate valuations of individual instruments, the risk factor scenarios are simulated so that portfolio value can be determined. Monte Carlo simulation uses historical data with more flexibility than the delta-gamma approach. After fitting a parametric model of the risk factors with stochastic volatility and return distributions preferably, one can simulate risk factor scenarios that can be defined in the model. The delta-gamma approach allows approximation of the risk factor dynamics, yet it uses partial valuation, which has a limited definition of risk factor scenarios based on the sensitivity parameters and the covariance matrix of historical data.

However, there are only a few samples drawn having $L > b$. When b is large, $L > b$ corresponds to "rare events." Importance sampling is a technique well suited for improving the efficiency of rare events simulation (Hammersley and Handscomb [1964]; Glynn and Iglehart [1989]; Glasserman [2004]). The basic idea of importance sampling is to replace the original density in Monte Carlo simulation by an alternative density for increasing the probability of rare event occurrence. For example, let X be a random variable with a density f. Given the goal of computing $P(X > 0)$, we want to calculate an expectation of the form $E\{I(X > 0)\}$ where $I(\cdot)$ is an indicator of the event of interests and E is the expectation associated with the original density f. Importance sampling involves choosing a sampling density g where there exists a likelihood ratio (or Radon-Nikodym derivative) ℓ such that

$$I(X > 0)f(X) = I(X > 0)\ell(X)g(X) \tag{5}$$

where

$$\ell(X) = \frac{f(X)}{g(X)}$$

We then have

$$P\big(X > 0\big) = E\big\{I(X > 0)\big\} = \tilde{E}\big\{I(X > 0)\ell(X)\big\}$$

where $\tilde{E}$ is the expectation associated with density g. One can then sample $I(X > 0)\ell(X)$ in the simulation from the importance sampling density g and estimate $P(X > 0)$ by the sample mean of $I(X > 0)\ell(X)$. The selection of the new density should be executed carefully, however, because simulation under an inappropriate density can increase estimator variance. In high dimension space, this problem is even more serious due to the variance of the likelihood ratio blowing up (Asmussen and Glynn [2007]). In implementing importance sampling, the key is to select an efficient importance sampling density g^* that ensures the importance sampling estimator has smaller variance (Asmussen and Glynn [2007]). Importance sampling has been studied in various applications (Glynn and Iglehart [1989]; Heidelberger [1995]; Glasserman [2004]). Regarding derivative valuation applications, Chiang et al. [2007], Chen and Glasserman [2008], and Joshi and Kainth [2004] have adopted importance sampling approaches to value basket default swaps.

Glasserman et al. [2000] developed stratified sampling and importance sampling variance reduction techniques for VaR. They assumed that the changes in the risk factors are normally distributed. Glasserman et al. [2002] extended the techniques to fit a heavy-tailed distribution. These techniques exploited the delta-gamma approach, which is not restricted by the assumption that portfolio losses move linearly with the changes in the risk factors. Hoogerheide et al. [2010] proposed an adaptive importance sampling method for forecasting VaR and ES in a Bayesian framework, which approximated the optimal importance sampling density by multi-step "high loss" scenarios.

The proposed algorithm calculates the accurate portfolio value at each time point, compared with the delta-gamma approximation. Also, the assumption of our algorithm is that the dependency structure of the risk factors is driven by a multivariate normal distribution. For example, the marginal distribution of each risk factor can be any distribution (Hull and White [1998]). Our approach focuses on the tail of loss distribution and simulates the excess (beyond threshold) portfolio losses directly without performing additional calculation for the complete loss distribution. In the case of estimating VaR and ES, simulation of the risk factors is also in a high dimension space. This article proposes a dimension reduction technique to ensure that the variance of the resulting likelihood ratios does not blow up.

The rest of this article is organized as follows. The research model being considered is defined and the assumption of the risk factors is established in the next section. An overview of the theoretical framework is then presented, which includes the central idea for identifying the key factor that drives the selection of importance sampling distribution and the implementation details of the proposed algorithm. The subsequent section presents the numerical results and a comprehensive comparison of the portfolios to demonstrate the efficiency of the proposed algorithm. The final section summarizes our main findings.

RESEARCH MODEL

Monte Carlo simulation relies on a suitable model of the risk factors. Thus, we first describe the dynamic behavior of the risk factors. Suppose there are d risk factors and $S(t)$ refers to the risk factors at time t. Let $S(t) = (S_1(t), \ldots, S_d(t))^T$ follow a multivariate diffusion process:[1]

$$\frac{dS(t)}{S(t)} = \mu(t)dt + \sigma(t)dZ(t) \tag{6}$$

where parameters $\mu(t) = (\mu_1(t),\ldots,\mu_d(t))$ and $\sigma(t) = (\sigma_1(t),\ldots,\sigma_d(t))$ are the drift and volatility of the process. The term $dZ(t)$ is a multivariate correlated Brownian motion with a correlation matrix ρ, where

$$\rho = \begin{pmatrix} 1 & \rho_{21} & \ldots & \ldots & \rho_{d1} \\ \rho_{12} & 1 & \ldots & \ldots & \rho_{d2} \\ \rho_{13} & \rho_{23} & 1 & \ldots & \rho_{d3} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ \rho_{1d} & \ldots & \ldots & \ldots & 1 \end{pmatrix}$$

Then, the covariance matrix of $dZ(t)$ for any $t > 0$ is

$$\begin{pmatrix} \sigma_1(t) & 0 & \ldots & \ldots & 0 \\ 0 & \sigma_2(t) & \ldots & \ldots & 0 \\ 0 & 0 & \sigma_3(t) & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & \ldots & \ldots & \ldots & \sigma_d(t) \end{pmatrix} \begin{pmatrix} 1 & \rho_{21} & \ldots & \ldots & \rho_{d1} \\ \rho_{12} & 1 & \ldots & \ldots & \rho_{d2} \\ \rho_{13} & \rho_{23} & 1 & \ldots & \rho_{d3} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ \rho_{1d} & \ldots & \ldots & \ldots & 1 \end{pmatrix} \begin{pmatrix} \sigma_1(t) & 0 & \ldots & \ldots & 0 \\ 0 & \sigma_2(t) & \ldots & \ldots & 0 \\ 0 & 0 & \sigma_3(t) & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & \ldots & \ldots & \ldots & \sigma_d(t) \end{pmatrix} dt \equiv \Sigma(t)dt$$

The log return of the risk factor S_i, is defined between time t and $t + \Delta t$ as X_i, where $i = 1, \ldots, d$. Let $X = (X_1, \ldots, X_d)^T$ and $a(t) = ((\mu_1(t) - \sigma_1^2(t)/2), \ldots, (\mu_d(t) - \sigma_d^2(t)/2))^T$. By the Itô lemma and the Euler scheme (Glasserman [2004]), it is straightforward to see that

$$X \sim N(a(t)\Delta t, \Sigma(t)\Delta t) \tag{7}$$

and

$$S_i(t + \Delta t) = S_i(t)e^{X_i}, \quad \text{where } i = 1, \ldots, d \tag{8}$$

Because $S(t + \Delta t)$ can be determined based on a given X, portfolio loss can be simulated by generating samples of X using Equation (1).

Monte Carlo simulation is a suitable approach for evaluating Equation (1). Values of the risk factors at time t are simulated, and the portfolio is reevaluated. Depending on the length of the horizon observed, distribution of the portfolio loss L over the given risk horizon Δt is produced. As previously mentioned, the problems of interest are $P(L > b)$ and $E[L; L > b]$. For a given loss threshold b, the estimator for $P(L > b)$ in Equation (3) is denoted by $\hat{p}$, as

$$\hat{p} = \frac{1}{n}\sum_{i=1}^{n} I(L_i > b)$$

and the standard error $(SE(\hat{p}))$ is $\sqrt{\hat{p}(1-\hat{p})/n}$. Similarly, the estimator for $E[L;\ L > b]$ in Equation (4) is denoted by $\hat{\gamma}$, as

$$\hat{\gamma} = \frac{\sum_{i=1}^{n} L_i I(L_i > b)}{n}$$

and its standard error $(SE(\hat{\gamma}))$ is $\sqrt{\sum_{i=1}^{n}(L_i I(L_i > b) - \hat{\gamma})^2/n(n-1)}$.

THE PROPOSED ALGORITHM

Two-Stage Sampling Procedure

The first goal is to determine the key factor from the covariance matrix of the risk factors. Hence, we begin by rewriting the log return of the risk factors in Equation (8) to

$$X = a(t)\Delta t + CZ \tag{9}$$

where C is a matrix satisfying $CC^T = \Sigma(t)\Delta t$ and $Z \sim N(0, I_d)$. One popular choice of C is the Cholesky factorization of $\Sigma(t)\Delta t$. Here, we use spectral decomposition. Let $(\lambda_1;\ q_1)$, ..., $(\lambda_d;\ q_d)$ be the eigen-pairs of $\Sigma(t)\Delta t$ and $\lambda_1 \geq \lambda_2, ..., \geq \lambda_d$. The spectral decomposition of $\Sigma(t)\Delta t$ is

$$\Sigma(t)\Delta t = \sum_{i=1}^{d} \lambda_i\, q_i\, q_i^T = Q\Lambda Q^T \tag{10}$$

and

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 & \cdots & \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_d \end{pmatrix}, \quad Q = (q_1, \ldots, q_d)$$

where Q is an orthogonal matrix and $\lambda_k \geq 0$. We set $C = Q\Lambda^{1/2}$, and it is clear that $Q\Lambda^{1/2}$ provides an alternative for C. Thus, $X = \alpha(t)\Delta t + Q\Lambda^{1/2}Z$ is a proper representation of X, and the risk factor at time $t + \Delta t$ is expressed as

$$S(t + \Delta t) = S(t)e^X = S(t)\exp\left[a(t)\Delta t + Q\Lambda^{1/2}Z)\right] \tag{11}$$

We decompose C as $(c_1, \tilde{C})$. The sub-matrixes c_1 and $\tilde{C}$ have the dimensions $d \times 1$ and $d \times (d - 1)$, respectively. Furthermore, the matrix Z is partitioned into $(Z_1, \tilde{Z})^T$. The sub-vectors Z_1 and $\tilde{Z}$ have the dimensions 1 and $d - 1$, respectively. Equation (11) can be rewritten as

$$S(t + \Delta t) = S(t)\exp\left[a(t)\Delta t + \left[\left(c_1 Z_1 + \tilde{C}\tilde{Z}\right)\right]\right] \tag{12}$$

Equation (12) provides the idea of designing a two-stage procedure to generate X. In particular, we first generate $\tilde{Z}$ and then Z_1. The relationship between Z_1 and the portfolio loss provides us a guideline for selecting a good importance sampling distribution. In our sampling procedure, an importance sampling distribution for X does not need to be defined. Instead, we generate $\tilde{Z}$ first, and then based on the sampled value of $\tilde{Z}$, we select importance sampling distributions for Z_1 given the sampling value of $\tilde{Z}$.

Based on Equation (12), it is clear that the portfolio loss is a function of Z_1 and $\tilde{Z}$. Set $h = L - b$, then h is also a function of Z_1 and $\tilde{Z}$. It is obvious that the event $L > b$ and $h > 0$ are equivalent.

Property 1. *The large loss probability $P(L > b)$ is equivalent to* $E[E[I(h > 0) \mid \tilde{Z}]]$.

Proof:

$$\begin{aligned} P(L > b) &= P(h > 0) \\ &= E[I(h > 0)] \\ &= E[E[I(h > 0) \mid \tilde{Z})]] \end{aligned}$$

That is, conditional on $\tilde{Z}$, the event of interest can be determined solely by Z_1. Our next goal is to determine the range of Z_1, such that $I(h > 0|\tilde{Z}) = 1$ for all $\tilde{Z}$. For achieving this goal, first we find a threshold z^* that makes $h = 0$ for a given $\tilde{Z}$. Second, we need to choose an appropriate new density for Z_1 that ensures the likelihood ratio is always less than one and forces the event of interest ($L > b$) to occur in every sample path. There are many possible candidates, however, for the new measures that can force the event of interest to occur. Hence, additional guidelines are necessary for determining an appropriate importance sampling distribution among all possible measures. For selecting a good density for Z_1, we need the specific monotonic property of $h(\cdot)$ in Z_1. It should be noted that the appropriate monotonic properties could be used for the guidelines of the selection of the new probability measure. The following lemmas provide such principles to put forward the new density.

Lemma 2. *For any $\tilde{Z}$, assume h is strictly monotonically decreasing in Z_1 and $h(z^*) = 0$. That is, for any $z < z^*$, $h(z) > 0$. Noted that z^* depends on the value of $\tilde{Z}$. We select the new density function $g(\cdot)$ for Z_1 as a truncated normal with a truncated region (z^*, ∞):*

$$g(z) = \begin{cases} \phi(z)/\Phi(z^*), & z < z^* \\ 0, & \text{otherwise} \end{cases}$$

where $\phi(\cdot)$ and $\Phi(\cdot)$ are pdf and cdf of a standard normal random variable, respectively.[2] *Thus, when $z < z^*$, the likelihood ratio of Z_1, $\ell = \Phi(z^*) < 1$, which is a function of $\tilde{Z}$.*

Proof: *For $z < z^*$*

$$\ell(z) = \frac{\phi(z)}{g(z)} = \frac{\phi(z)}{\frac{\phi(z)}{\Phi(z^*)}} = \Phi(z^*)$$

Lemma 3. *For any $\tilde{Z}$, assume h is strictly monotonically increasing in Z_1 and $h(z^{**}) = 0$. That is, for any $z > z^{**}$, $h(z) > 0$. Noted that z^{**} depends on the value of $\tilde{Z}$. We select the new density function $g(\cdot)$ for Z_1 as a truncated normal with a truncated region $(-\infty, z^{**})$:*

$$g(z) = \begin{cases} \phi(z)/(1 - \Phi(z^{**})), & z > z^{**} \\ 0, & \text{otherwise} \end{cases}$$

Thus, when $z > z^{**}$ the likelihood ratio of Z_1, $\ell = 1 - \Phi(z^{**}) < 1$, which is a function of $\tilde{Z}$.

Proof: *For $z > z^{**}$*

$$\ell(z) = \frac{\phi(z)}{g(z)} = \frac{\phi(z)}{\frac{\phi(z)}{1 - \Phi(z^{**})}} = 1 - \Phi(z^{**})$$

Lemma 4. *For any $\tilde{Z}$, assume h is a convex function in Z_1 and $z^* < z^{**}$ are the roots of h, then for any $z < z^*$ or $z > z^{**}$, $h(z) > 0$. Noted that (z^*, z^{**}) depends on the value of $\tilde{Z}$. We select the new density function $g(\cdot)$ for Z_1 as a truncated normal with a truncated region (z^*, z^{**}):*

$$g(z) = \begin{cases} \phi(z) / (\Phi(z^{**}) + (1 - \Phi(z^{**})), & z < z^* \text{ or } z > z^{**} \\ 0, & \text{otherwise} \end{cases}$$

Thus, when $z < z^*$ or $z > z^{**}$, the likelihood ratio of Z_1, $\ell = \Phi(z^*) + (1 - \Phi(z^{**})) < 1$, *which is a function of $\tilde{Z}$.*

Proof: *For* $z < z^*$ *or* $z > z^{**}$,

$$\ell(z) = \frac{\phi(z)}{g(z)} = \frac{\phi(z)}{\frac{\phi(z)}{\Phi(z^*) + (1 - \Phi(z^{**}))}} = \Phi(z^*) + (1 - \Phi(z^{**}))$$

Briefly, these lemmas provide us a guideline to select an appropriate importance sampling distribution. Let $\ell_1, \ldots, \ell_n$ be independent copies of ℓ, and then we can derive the importance sampling estimator for $P(L > b)$ as

$$\hat{p}_{IS} = \frac{1}{n}\sum_{i=1}^{n} \ell_i I(L > b) = \frac{1}{n}\sum_{i=1}^{n} \ell_i \tag{13}$$

and the standard error can be estimated by $\sqrt{\sum_{i=1}^{n}(\ell_i - \hat{p}_{IS})^2/n(n-1)}$. Furthermore, we now consider the application of this algorithm to estimate $E[L; L > b]$. Define

$$\hat{\gamma}_{IS} = E_{IS}[L; L > b] = \frac{1}{n}\sum_{i=1}^{n} \tilde{L}_i \ell_i \tag{14}$$

and

$$\tilde{L}_i = V(S_i(t)) - V(\tilde{S}_i(t + \Delta t)) \equiv f(S_i(t), \tilde{S}_i(t + \Delta t))$$

where $\tilde{L}_i$ is the simulated loss and $\tilde{S}(t)$ is the simulated risk factors under the new density. Note that this algorithm does not require the risk factors to be restricted to be normally distributed, as long as the dependency structure of the risk factors is driven by a normal distribution (Hull and White [1998]). In addition to the benefits listed, the advantage of our approach is the variance reduction established according to the following proposition.

Proposition 5. *Since* $\mathrm{P}(L > b) = E[I(L > b)] = E[I(h > 0)]$, *based on our importance sampling density, the likelihood ratio* $\ell(\cdot) = E[I(L > b) \mid \tilde{Z}]$. *Its variance is smaller than the variance of* $I(L > b)$ *by conditional expectation properties (Asmussen and Glynn [2007]).*

Proof:

$$Var(I(L > b)) > Var(E[I(h > 0) \mid \tilde{Z}]) = Var(\ell(\cdot))$$

Proposition 5 proves that the proposed algorithm guarantees reduction in variance compared with naive Monte Carlo simulation.

Importance Sampling Algorithm

The previous analysis establishes the theoretical background for computing the new measure. In the following, we present the algorithm for estimating $P(L > b)$ and $E[L;\ L > b]$, respectively. For the simple and straightforward idea, we assume the portfolio loss is monotonically decreasing in Z_1.

Goal*: Estimate* $P(L > b)$ *and* $E[L;\ L > b]$
Algorithm:

1. Initialization
   a. Compute $V(t) = V(S(t))$
   b. Compute Q and Λ by spectral decomposition, where $\sum(t)\Delta t = Q\Lambda Q^T$
   c. Express $C = Q\Lambda^{1/2}$
   d. Partition C into sub-matrixes $(c_1, \tilde{C})$ of size $d \times 1$ and $d \times (d - 1)$; where $c_1 = q_1\lambda_1^{1/2}$ and (λ_1, q_1) is the largest eigenvalue and its eigenvector.
   e. Express the portfolio terminal price as

$$V(t+\Delta t) = V(S(t+\Delta t)) = S(t)e^X = S(t)e^{(\mu(t)-\sigma^2/2)\Delta t + c_1 Z_1 + \tilde{C}\tilde{Z}}$$

2. Perform the importance sampling simulation: For each $i = 1, \ldots, n$
   a. Execute the proposed two-stage sampling procedure
      i. First stage:
         1) Draw $\tilde{Z} \sim N(0, I_{d-1})$ and generate $\tilde{C}\tilde{Z}$ for every replication.
      ii. Second stage:
         1) Set the large loss event as

$$L > b = V(S(t)) - V\left(S(t)e^{a(t)\Delta t + \left[\left(c_1 Z_1 + \tilde{C}\tilde{Z}\right)\right]}\right) - b > 0 = h > 0$$

         2) Solve $z^{*(i)}$ such that $h(z^{*(i)}) = 0$
   b. Generate the likelihood ratio $\ell_i = \Phi\left(z^{*(i)}\right)$
   c. Draw $Z^{IS(i)} = \Phi^{-1}(U_i\ z^{*(i)})$, where $U_i \sim$ uniform (0, 1)
   d. Using $Z^{IS(i)}$, evaluate simulated loss, $\tilde{L}_i = V(S_i(t)) - V(\tilde{S}_i(t+\Delta t))$, where $V(\tilde{S}_i(t+\Delta t)) = V\left(S_i(t)e^{a\Delta t + \left[\left(c_1 Z^{IS(i)} + \tilde{C}\tilde{Z}\right)\right]}\right)$, and $\tilde{S}(t)$ are simulated terminal risk factor price using simulated $Z^{IS(i)}$.
3. Report results
   a. Estimate $P(L > b)$ by $\frac{1}{n}\Sigma_{i=1}^{n}\ell_i$
   b. Estimate $E[L;\ L > b]$ by $\frac{1}{n}\Sigma_{i=1}^{n}\tilde{L}_i\ell_i$.

Note that if the portfolio loss is monotonically increasing in Z_1, we only have to replace the likelihood ratio to $\ell_i = 1 - \Phi(z^{**(i)})$. Similarly, for the case that the portfolios loss is a convex function in Z_1, the likelihood ratio should adopt $\ell_i = \Phi(z^{*(i)}) + (1 - \Phi(z^{**(i)}))$.

NUMERICAL EXAMPLES

In this article, the test portfolios for the numerical examples follow Glasserman et al. [2000] but replace the uncorrelated normal assumptions on the risk factors of that paper with the correlated Gaussian assumptions of this study. The subsets of the test portfolios are option portfolios that consist of European call options and European put options. The nonlinear price characteristic of option positions implies the difficulty of estimating portfolio loss, and the distribution of the portfolio loss is more complex than linear portfolios. Exhibit 1 describes the test portfolios. We value the options using the Black–Scholes formula because it highlights the nonlinear price characteristic of options. The first is an option portfolio with short call and short put in portfolio (a.1) (see Exhibit 1). The second is a perfectly delta-neutral portfolio in (b.1). Next, we investigate the effect of the moderate correlation in portfolios (a.2) and (a.3) and high correlation in (b.2) and (b.3).

EXHIBIT 1

Test Portfolios for the Numerical Example

(a.1) Index	Short 50 at-the-money calls and 50 at-the-money puts on 10 underlying risk factors and all options have a maturity of 0.5 years. The covariance matrix for 10 asset prices is set to be the same as that in Glasserman et al. [2000] in Exhibit 3. The initial risk factors prices are set as (100, 50, 20, 100, 80, 20, 50, 200, 150, 10). The eigenvalues $\boldsymbol{\lambda}$ of the covariance matrix are (0.729, 0.226, 0.112, 0.060, 0.045, 0.032, 0.024, 0.020, 0.015, 0.012).
(a.2) Moderate Correlated, ρ=0.5	Same as (a.1), but set the risk factors as moderately correlated. The covariance matrix for 10 asset prices is as Exhibit 4. The eigenvalues λ of the covariance matrix are (0.723, 0.126, 0.085, 0.077, 0.072, 0.061, 0.045, 0.039, 0.035, 0.012).
(a.3) Highly Correlated, ρ=0.9	Same as (a.1). Set the risk factors as highly correlated. The covariance matrix for 10 asset prices is as Exhibit 5. The eigenvalues λ of the covariance matrix are (1.163, 0.026, 0.017, 0.015, 0.014, 0.012, 0.009, 0.008, 0.007, 0.002).
(b.1) Delta-hedged	Short 10 at-the-money calls on each risk factor; all options have a maturity of 0.1 years but with the numbers of puts on each risk factor increasing so that delta becomes 0, i.e., the portfolio is delta-neutral. The initial asset prices and covariance matrix are set as in case (a.1).
(b.2) Delta-hedged, ρ=0.5	Same as (b.1), but set the risk factors as moderately correlated, and the covariance matrix of the risk factors is as Exhibit 4.
(b.3) Delta-hedged, ρ=0.9	Same as (b.1). Set the risk factors as highly correlated. The covariance matrix of the risk factors is as Exhibit 5.

We assume 250 trading days in a year and a continuously compounded risk-free rate of interest of 5%. For each case, we investigate losses over one day ($\Delta t = 0.004$ years). Exhibit 2 lists the given parameters in all experiments.

Exhibit 3 shows the covariance matrix from Glasserman et al. [2000]. Exhibits 4 and 5 are the covariance matrixes for $\rho = 0.5$ and $\rho = 0.9$, respectively. The volatility of the assets is the same for all portfolios. Set the volatilities of the assets as (0.538, 0.341, 0.148, 0.281, 0.259, 0.383, 0.397, 0.277, 0.378, 0.420).

We compare two methods—naive Monte Carlo (MC) and importance sampling (IS). In each case, we generate 1,000,000 simulation runs for the naive Monte Carlo approach but only 5,000 for the importance sampling method. We define the variance reduction ratio (VRR) to quantify the statistical efficiency of the importance sampling estimator as follows:

$$VRR = \frac{\text{Sample size of naive MC estimator} * Var(\text{Native MC estimator})}{\text{Sample size of IS estimator} * Var(\text{IS estimator})} \tag{15}$$

Numerical Results for the Large Loss Probability

For meeting the requirements of market risk when calculating daily VaR, we first adjust the loss threshold b so that the probability to be estimated is close to 0.05 and 0.01 and then move to extreme losses. In Exhibit 6, we report loss threshold b in column 2, the point estimators in columns 3 and 4, the standard errors in columns 5 and 6, and the coefficients of variation in columns 7 and 8 for each test portfolio. The results indicate that the importance sampling estimator is very close to the

EXHIBIT 2
Parameter Settings

		(a.1),(a.2),(a.3)	*(b.1),(b.2),(b.3)*
R	Short rate	0.05	0.05
T	Time to maturity	0.5	0.1
Δt	Risk horizon	0.004	0.004

EXHIBIT 3
Covariance Matrix for Portfolios (a.1) and (b.1)

$$\begin{pmatrix} 0.289 & 0.069 & 0.008 & 0.069 & 0.084 & 0.085 & 0.081 & 0.052 & 0.075 & 0.114 \\ 0.069 & 0.116 & 0.020 & 0.061 & 0.036 & 0.088 & 0.102 & 0.070 & 0.005 & 0.102 \\ 0.008 & 0.020 & 0.022 & 0.013 & 0.009 & 0.016 & 0.019 & 0.016 & 0.010 & 0.017 \\ 0.069 & 0.061 & 0.013 & 0.079 & 0.035 & 0.090 & 0.090 & 0.051 & 0.031 & 0.075 \\ 0.084 & 0.036 & 0.009 & 0.035 & 0.067 & 0.055 & 0.049 & 0.029 & 0.022 & 0.062 \\ 0.085 & 0.088 & 0.016 & 0.090 & 0.055 & 0.147 & 0.125 & 0.073 & 0.016 & 0.112 \\ 0.081 & 0.102 & 0.019 & 0.090 & 0.049 & 0.125 & 0.158 & 0.087 & 0.016 & 0.127 \\ 0.052 & 0.070 & 0.016 & 0.051 & 0.029 & 0.073 & 0.087 & 0.077 & 0.014 & 0.084 \\ 0.075 & 0.005 & 0.010 & 0.031 & 0.022 & 0.016 & 0.016 & 0.014 & 0.143 & 0.033 \\ 0.114 & 0.102 & 0.017 & 0.075 & 0.062 & 0.112 & 0.127 & 0.084 & 0.033 & 0.176 \end{pmatrix}$$

EXHIBIT 4
Covariance Matrix for Portfolios (a.2) and (b.2)

$$\begin{pmatrix}
0.289 & 0.092 & 0.040 & 0.076 & 0.070 & 0.103 & 0.107 & 0.075 & 0.102 & 0.113 \\
0.092 & 0.116 & 0.025 & 0.048 & 0.044 & 0.065 & 0.068 & 0.047 & 0.064 & 0.071 \\
0.040 & 0.025 & 0.022 & 0.021 & 0.019 & 0.028 & 0.029 & 0.021 & 0.028 & 0.031 \\
0.076 & 0.048 & 0.021 & 0.079 & 0.036 & 0.054 & 0.056 & 0.039 & 0.053 & 0.059 \\
0.070 & 0.044 & 0.019 & 0.036 & 0.067 & 0.050 & 0.051 & 0.036 & 0.049 & 0.054 \\
0.103 & 0.065 & 0.028 & 0.054 & 0.050 & 0.147 & 0.076 & 0.053 & 0.072 & 0.080 \\
0.107 & 0.068 & 0.029 & 0.056 & 0.051 & 0.076 & 0.158 & 0.055 & 0.075 & 0.083 \\
0.075 & 0.047 & 0.021 & 0.039 & 0.036 & 0.053 & 0.055 & 0.077 & 0.052 & 0.058 \\
0.102 & 0.064 & 0.028 & 0.053 & 0.049 & 0.072 & 0.075 & 0.052 & 0.143 & 0.079 \\
0.113 & 0.071 & 0.031 & 0.059 & 0.054 & 0.080 & 0.083 & 0.058 & 0.079 & 0.176
\end{pmatrix}$$

EXHIBIT 5
Covariance Matrix for Portfolios (a.3) and (b.3)

$$\begin{pmatrix}
0.289 & 0.165 & 0.072 & 0.136 & 0.125 & 0.186 & 0.192 & 0.134 & 0.183 & 0.203 \\
0.165 & 0.116 & 0.045 & 0.086 & 0.079 & 0.118 & 0.122 & 0.085 & 0.116 & 0.129 \\
0.072 & 0.045 & 0.022 & 0.038 & 0.035 & 0.051 & 0.053 & 0.037 & 0.050 & 0.056 \\
0.136 & 0.086 & 0.038 & 0.079 & 0.065 & 0.097 & 0.101 & 0.070 & 0.096 & 0.106 \\
0.125 & 0.079 & 0.035 & 0.065 & 0.067 & 0.089 & 0.093 & 0.065 & 0.088 & 0.098 \\
0.186 & 0.118 & 0.051 & 0.097 & 0.089 & 0.147 & 0.137 & 0.096 & 0.130 & 0.145 \\
0.192 & 0.122 & 0.053 & 0.101 & 0.093 & 0.137 & 0.158 & 0.099 & 0.135 & 0.150 \\
0.134 & 0.085 & 0.037 & 0.070 & 0.065 & 0.096 & 0.099 & 0.077 & 0.094 & 0.105 \\
0.183 & 0.116 & 0.050 & 0.096 & 0.088 & 0.130 & 0.135 & 0.094 & 0.143 & 0.143 \\
0.203 & 0.129 & 0.056 & 0.106 & 0.098 & 0.145 & 0.150 & 0.105 & 0.143 & 0.176
\end{pmatrix}$$

naive Monte Carlo simulation. The variance reduction ratios for 1% $\hat{p}$ are apparently more efficient than 5% $\hat{p}$. We find the correlation between the risk factors has strong impacts on the order of the variance improvements. For example, the variance improvement is about 10^2 in (a.2) and increases one order to 10^3 in portfolio (a.3). Furthermore, the importance sampling estimator is more effective in delta-hedged portfolios than not delta-hedged. Dramatic variance reduction is obtained when the risk factors are highly correlated (i.e., VRR is almost 10^5 in portfolio (b.3)).

Next, we move the loss threshold b to extreme high levels in Exhibits 7 to 12. We find the method performs better for higher loss levels; for example, the variance improvements are 261, 3.765×10^3, and 1.163×10^4, when the loss threshold moves, respectively, to 400, 600, and 800 in portfolio (a.1).

EXHIBIT 6
Results for 5% $\hat{p}_{IS}$ and 1% $\hat{p}_{IS}$ of All Portfolios

Portfolio	α	b	$\hat{p}$	$\hat{p}_{IS}$	$SE(\hat{p})$	$SE(\hat{p}_{IS})$	$CV(\hat{\gamma})$	$CV(\hat{\gamma}_{IS})$	VRR
(a.1)	5%	210	0.05	0.051	0.00022	0.00048	4.348	0.677	41
	1%	329	0.01	0.01	0.00010	0.00013	9.949	0.902	119
(a.2)	5%	228	0.05	0.05	0.000218	0.000368	4.352	0.516	70
	1%	360	0.0098	0.0099	0.000099	0.000098	10.038	0.693	202
(a.3)	5%	297	0.0496	0.0498	0.000217	0.000118	4.374	0.167	677
	1%	477	0.0099	0.010	0.000099	0.000031	9.967	0.217	2073
(b.1)	5%	24	0.0499	0.04773	0.002217	0.000669	4.364	0.991	21
	1%	44	0.01	0.00956	0.000099	0.000113	9.948	0.839	154
(b.2)	5%	24	0.0538	0.0534	0.000226	0.000359	4.192	0.475	78
	1%	48	0.0104	0.0102	0.000102	0.000073	9.733	0.505	383
(b.3)	5%	38	0.05	0.051	0.000220	0.000030	4.318	0.041	10580
	1%	75	0.0101	0.0102	0.000100	0.000007	9.943	0.051	36460

EXHIBIT 7
Results for $\hat{p}_{IS}$ of Portfolio (a.1)

Loss	$\hat{p}$	$\hat{p}_{IS}$	$SE(\hat{p})$	$SE(\hat{p}_{IS})$	$CV(\hat{\gamma})$	$CV(\hat{\gamma}_{IS})$	VRR
400	$3.529*10^{-3}$	$3.532*10^{-3}$	$5.930*10^{-5}$	$5.189*10^{-5}$	$1.680*10$	1.039	261
500	$7.610*10^{-4}$	$7.338*10^{-4}$	$2.758*10^{-5}$	$1.281*10^{-5}$	$3.624*10$	1.235	926
600	$1.510*10^{-4}$	$1.395*10^{-4}$	$1.229*10^{-5}$	$2.832*10^{-6}$	$8.137*10$	1.436	$3.765*10^{3}$
700	$1.900*10^{-5}$	$2.461*10^{-5}$	$4.359*10^{-5}$	$5.714*10^{-7}$	$2.294*10^{2}$	1.642	$1.163*10^{4}$
800	$7.000*10^{-6}$	$4.076*10^{-6}$	$2.646*10^{-6}$	$1.068*10^{-7}$	$3.780*10^{2}$	1.854	$1.226*10^{5}$

We also compare the efficiency of the estimators at the same level, about 10^{-4} to all portfolios (i.e., $b = 500$ in (a.1); $b = 600$ in (a.2); $b = 900$ in (a.3); $b = 120$ in (b.1); $b = 110$ in (b.2); $b = 180$ in (b.3)). Our algorithm results in a significant decrease in variance for the risk factors that are highly correlated. For example, the variance reduction is 9.26×10^2 in (a.1), 2.603×10^3 in (a.2), and 7.483×10^4 in (a.3). With respect to the delta-hedged portfolios, our algorithm performs better, and the variance improvement is about one order more than not

EXHIBIT 8
Results for $\hat{p}_{IS}$ of Portfolio (a.2)

Loss	$\hat{p}$	$\hat{p}_{IS}$	$SE(\hat{p})$	$SE(\hat{p}_{IS})$	$CV(\hat{\gamma})$	$CV(\hat{\gamma}_{IS})$	VRR
400	$5.709*10^{-3}$	$5.902*10^{-3}$	$7.534*10^{-5}$	$6.241*10^{-5}$	$1.320*10$	$7.478*10^{-1}$	291
450	$2.893*10^{-3}$	$2.993*10^{-3}$	$5.371*10^{-5}$	$3.454*10^{-5}$	$1.857*10$	$8.159*10^{-1}$	483
500	$1.423*10^{-3}$	$1.487*10^{-3}$	$3.770*10^{-5}$	$1.860*10^{-5}$	$2.649*10$	$8.845*10^{-1}$	821
550	$7.060*10^{-4}$	$7.254*10^{-4}$	$2.656*10^{-5}$	$9.783*10^{-6}$	$3.762*10$	$9.536*10^{-1}$	$1.474*10^{3}$
600	$3.300*10^{-4}$	$3.478*10^{-4}$	$1.816*10^{-5}$	$5.034*10^{-6}$	$5.504*10$	1.023	$2.603*10^{3}$

EXHIBIT 9
Results for $\hat{p}_{IS}$ of Portfolio (a.3)

Loss	$\hat{p}$	$\hat{p}_{IS}$	$SE(\hat{p})$	$SE(\hat{p}_{IS})$	$CV(\hat{\gamma})$	$CV(\hat{\gamma}_{IS})$	VRR
600	$3.085*10^{-3}$	$3.099*10^{-3}$	$5.546*10^{-5}$	$1.089*10^{-5}$	$1.798*10$	$2.484*10^{-1}$	$5.19*10^{3}$
700	$1.114*10^{-3}$	$1.145*10^{-3}$	$3.336*10^{-5}$	$4.414*10^{-6}$	$2.994*10$	$2.727*10^{-1}$	$1.142*10^{4}$
800	$4.110*10^{-4}$	$4.103*10^{-4}$	$2.027*10^{-5}$	$1.718*10^{-6}$	$4.932*10$	$2.961*10^{-1}$	$2.784*10^{4}$
900	$1.550*10^{-4}$	$1.433*10^{-4}$	$1.245*10^{-5}$	$6.455*10^{-7}$	$8.032*10$	$3.186*10^{-1}$	$7.438*10^{4}$
1000	$5.400*10^{-5}$	$4.888*10^{-5}$	$7.348*10^{-6}$	$2.353*10^{-7}$	$1.361*10^{2}$	$3.405*10^{-1}$	$1.95*10^{5}$

EXHIBIT 10
Results for $\hat{p}_{IS}$ of Portfolio (b.1)

Loss	$\hat{p}$	$\hat{p}_{IS}$	$SE(\hat{p})$	$SE(\hat{p}_{IS})$	$CV(\hat{\gamma})$	$CV(\hat{\gamma}_{IS})$	VRR
60	$6.324*10^{-3}$	$6.040*10^{-3}$	$7.927*10^{-5}$	$7.070*10^{-5}$	$1.254*10$	$8.277*10^{-1}$	251
90	$6.670*10^{-4}$	$6.337*10^{-4}$	$2.582*10^{-5}$	$7.759*10^{-6}$	$3.871*10$	$8.658*10^{-1}$	$2.214*10^{3}$
120	$3.230*10^{-4}$	$3.012*10^{-4}$	$1.797*10^{-5}$	$3.821*10^{-6}$	$5.563*10$	$8.970*10^{-1}$	$4.423*10^{3}$
150	$4.400*10^{-5}$	$3.260*10^{-5}$	$6.633*10^{-6}$	$4.728*10^{-7}$	$1.508*10^{2}$	1.025	$3.937*10^{4}$
180	$4.000*10^{-6}$	$3.536*10^{-6}$	$2.000*10^{-6}$	$5.985*10^{-8}$	$5.000*10^{2}$	1.197	$2.2337*10^{5}$

EXHIBIT 11
Results for $\hat{p}_{IS}$ of Portfolio (b.2)

Loss	$\hat{p}$	$\hat{p}_{IS}$	$SE(\hat{p})$	$SE(\hat{p}_{IS})$	$CV(\hat{\gamma})$	$CV(\hat{\gamma}_{IS})$	VRR
70	$2.448*10^{-3}$	$2.391*10^{-3}$	$4.942*10^{-5}$	$2.104*10^{-5}$	$2.019*10$	$6.224*10^{-1}$	$1.103*10^{3}$
80	$1.320*10^{-4}$	$1.242*10^{-3}$	$3.631*10^{-5}$	$1.202*10^{-5}$	$2.751*10$	$6.843*10^{-1}$	$1.826*10^{3}$
90	$6.980*10^{-4}$	$6.469*10^{-4}$	$2.641*10^{-5}$	$6.871*10^{-6}$	$3.784*10$	$7.511*10^{-1}$	$2.955*10^{3}$
100	$3.820*10^{-4}$	$3.377*10^{-4}$	$1.954*10^{-5}$	$3.929*10^{-6}$	$5.115*10$	$8.227*10^{-1}$	$4.948*10^{3}$
110	$2.020*10^{-4}$	$1.765*10^{-4}$	$1.421*10^{-5}$	$2.245*10^{-6}$	$7.035*10$	$8.993*10^{-1}$	$8.014*10^{3}$

EXHIBIT 12
Results for $\hat{p}_{IS}$ of Portfolio (b.3)

Loss	$\hat{p}$	$\hat{p}_{IS}$	$SE(\hat{p})$	$SE(\hat{p}_{IS})$	$CV(\hat{\gamma})$	$CV(\hat{\gamma}_{IS})$	*VRR*
120	$1.457*10^{-3}$	$1.519*10^{-3}$	$3.814*10^{-5}$	$1.507*10^{-6}$	$7.612*10$	$6.497*10^{-1}$	$1.28*10^{5}$
140	$6.460*10^{-4}$	$6.553*10^{-4}$	$2.541*10^{-5}$	$7.545*10^{-7}$	$1.274*10^{2}$	$7.117*10^{-1}$	$2.268*10^{5}$
160	$2.730*10^{-4}$	$2.830*10^{-4}$	$1.652*10^{-5}$	$3.778*10^{-7}$	$2.192*10^{2}$	$7.612*10^{-1}$	$3.824*10^{5}$
180	$1.220*10^{-4}$	$1.222*10^{-4}$	$1.104*10^{-5}$	$1.886*10^{-7}$	$3.707*10^{2}$	$8.090*10^{-1}$	$6.860*10^{5}$
200	$5.500*10^{-5}$	$5.272*10^{-5}$	$7.416*10^{-6}$	$9.363*10^{-8}$	$6.396*10^{2}$	$9.207*10^{-1}$	$1.254*10^{6}$

EXHIBIT 13
Results for 5% $\hat{\gamma}_{IS}$ and 1% $\hat{\gamma}_{IS}$ of All Portfolios

Portfolio	α	*Loss*	$\hat{\gamma}$	$\hat{\gamma}_{IS}$	$SE(\hat{\gamma})$	$SE(\hat{\gamma}_{IS})$	$CV(\hat{\gamma})$	$CV(\hat{\gamma}_{IS})$	*VRR*
(a.1)	5%	210	14.234	14.397	0.069	0.152	4.889	0.747	41
	1%	329	3.962	3.986	0.064	0.052	16.331	0.940	298.
(a.2)	5%	228	15.461	15.476	0.076	0.134	4.917	0.613	64
	1%	360	4.253	4.328	0.070	0.044	16.536	0.733	491
(a.3)	5%	307	20.301	20.272	0.106	0.089	5.25	0.311	285
	1%	500	5.786	5.841	0.100	0.023	17.359	0.279	3790
(b.1)	5%	23	1.808	2.053	0.0131	0.137	7.257	4.745	1.8
	1%	44	0.573	0.55	0.0135	0.008	23.641	1.058	540
(b.2)	5%	24	2.081	2.082	0.015	0.059	7.27	2.023	13
	1%	48	0.662	0.646	0.015	0.005	23.751	0.586	1720
(b.3)	5%	38	3.096	3.111	0.023	0.016	7.476	0.38	383
	1%	75	0.984	1.001	0.023	0.003	24.022	0.246	9236

delta-hedged. For example, the VRR is 4.423×10^3 in (b.1), 8.014×10^3 in (b.2), and 3.824×10^5 in (b.3) when the $\hat{p}$ is about 10^{-4}.

Furthermore, we see that the coefficient of variation is nearly constant across a broad range of b values when estimating $\hat{p}_{IS}$. This indicates that the importance sampling estimator has a bounded relative error. Estimators with bounded relative error are best-of-class importance sampling estimators (Heidelberger [1995]).

Finally, we examine the effect of the eigenvalues. We find that if the largest eigenvalue strongly dominates the others (i.e., (a.3) and (b.3)), the variance decreases more. The performance seems to be associated with the size of the largest eigenvalue. However, our method demonstrates accurate estimations under all conditions not influenced by the size of the largest eigenvalue. Even if the first two largest eigenvalues are identical, our method still gains a major improvement.

Numerical Results of the Expected Exceeded Loss Above Threshold

Results for $\hat{\gamma}_{IS}$ are given in Exhibit 13. Our algorithm obtains similar outcomes as experiments with $\hat{p}_{IS}$. High correlation between the risk factors guarantees large variance reduction (i.e., comparing (a.3) and (b.3) with (a.1) and (b.1)). In addition, leaving out (b.3), by comparing Exhibit 6 with Exhibit 13 at the same level of the loss threshold, the performance of $\hat{\gamma}_{IS}$ seems to

EXHIBIT 14
Results for $\hat{\gamma}_{IS}$ of Portfolio (a.1)

Loss	$\hat{\gamma}$	$\hat{\gamma}_{IS}$	$SE(\hat{\gamma})$	$SE(\hat{\gamma}_{IS})$	$CV(\hat{\gamma})$	$CV(\hat{\gamma}_{IS})$	*VRR*
400	1.640	1.639	$6.213*10^{-2}$	$2.499*10^{-2}$	$3.787*10$	1.078	$1.236*10^{3}$
500	$4.270*10^{-1}$	$4.122*10^{-1}$	$5.860*10^{-2}$	$7.471*10^{-3}$	$1.372*10^{2}$	1.282	$1.23*10^{4}$
600	$9.892*10^{-2}$	$9.161*10^{-2}$	$5.683*10^{-2}$	$1.876*10^{-3}$	$5.745*10^{2}$	1.448	$1.835*10^{5}$
700	$1.477*10^{-2}$	$1.858*10^{-2}$	$6.051*10^{-2}$	$4.337*10^{-4}$	$4.098*10^{3}$	1.651	$3.891*10^{6}$
800	$5.930*10^{-3}$	$3.493*10^{-3}$	$1.859*10^{-2}$	$9.447*10^{-5}$	$3.134*10^{3}$	1.912	$7.741*10^{6}$

EXHIBIT 15
Results for $\hat{\gamma}_{IS}$ of Portfolio (a.2)

Loss	$\hat{\gamma}$	$\hat{\gamma}_{IS}$	$SE(\hat{\gamma})$	$SE(\hat{\gamma}_{IS})$	$CV(\hat{\gamma})$	$CV(\hat{\gamma}_{IS})$	*VRR*
400	2.694	2.790	$6.904*10^{-2}$	$3.106*10^{-2}$	$2.562*10$	$7.870*10^{-1}$	988
450	1.504	1.557	$6.741*10^{-2}$	$1.859*10^{-2}$	$4.482*10$	$8.445*10^{-1}$	$2.628*10^{3}$
500	$8.088*10^{-1}$	$8.421*10^{-1}$	$6.636*10^{-2}$	$1.076*10^{-2}$	$8.205*10$	$9.035*10^{-1}$	$7.606*10^{3}$
550	$4.342*10^{-1}$	$4.479*10^{-1}$	$6.587*10^{-2}$	$6.197*10^{-3}$	$1.517*10^{2}$	$9.783*10^{-1}$	$2.259*10^{4}$
600	$2.192*10^{-1}$	$2.309*10^{-1}$	$6.723*10^{-2}$	$3.390*10^{-3}$	$3.067*10^{2}$	1.038	$7.865*10^{4}$

EXHIBIT 16
Results for $\hat{\gamma}_{IS}$ of Portfolio (a.3)

Loss	$\hat{\gamma}$	$\hat{\gamma}_{IS}$	$SE(\hat{\gamma})$	$SE(\hat{\gamma}_{IS})$	$CV(\hat{\gamma})$	$CV(\hat{\gamma}_{IS})$	*VRR*
600	2.158	2.169	$9.763*10^{-2}$	$8.724*10^{-3}$	$4.523*10$	$2.845*10^{-1}$	$2.504*10^{4}$
700	$8.913*10^{-1}$	$9.118*10^{-1}$	$9.585*10^{-2}$	$3.876*10^{-3}$	$1.075*10^{2}$	$3.006*10^{-1}$	$1.222*10^{5}$
800	$3.694*10^{-1}$	$3.667*10^{-1}$	$9.010*10^{-2}$	$1.640*10^{-3}$	$2.439*10^{2}$	$3.163*10^{-1}$	$6.034*10^{5}$
900	$1.534*10^{-1}$	$1.418*10^{-1}$	$8.303*10^{-2}$	$6.692*10^{-4}$	$5.412*10^{2}$	$3.337*10^{-1}$	$3.079*10^{6}$
1000	$5.829*10^{-2}$	$5.326*10^{-2}$	$7.745*10^{-2}$	$2.671*10^{-4}$	$1.329*10^{3}$	$3.546*10^{-1}$	$1.681*10^{7}$

be slightly better than $\hat{p}_{IS}$ (i.e., for 1% $\hat{p}_{IS}$ in Exhibit 6 versus 1% $\hat{\gamma}_{IS}$ in Exhibit 13, the VRR is 383 and 1,720 in (b.2)).

Results for extreme losses are given in Exhibits 14 to 19. The variance improvements in extreme losses result in significant efficiency (i.e., the VRR in (a.1) is about 10^6 for $b = 800$ in Exhibit 14). In addition, the boost in performance is increasing a lot for extreme b. For example, the variances decreases are about 10^5, 10^6, and 10^7 for portfolio (a.3) where b moves to 800, 900, and 1,000, respectively.

Furthermore, by comparing the same level of estimator that is close to 0.4 (i.e., $b = 500$ in (a.1); $b = 550$ in (a.2); $b = 800$ in (a.3); $b = 60$ in (b.1); $b = 70$ in (b.2); $b = 120$ in (b.3)), our algorithm results in a significant decrease in variance for a portfolio with highly correlated risk factors. However, there seems to be a slight increase in variance for delta-hedged portfolios. For example, the VRR is 1.236×10^3 in (a.1), 2.259×10^4 in (a.2), and 6.034×10^5 in (a.3), but the VRR decreases slightly to 1.162×10^3 in (b.1), 1.645×10^4 in (b.2), and 3.569×10^5 in (b.3).

Conclusively, the standard error of $\hat{\gamma}_{IS}$ is decreasing in all portfolios, and the coefficient of variation is always constant as $\hat{p}_{IS}$. These characteristics demonstrate that our algorithm has a bounded relative error and contribute with a significant variance reduction ratio in a stable manner.

CONCLUSION

This article proposes a fast Monte Carlo algorithm using importance sampling technique for estimation of large loss probability, $\hat{p}_{IS}(L > b)$, and the expected loss above the threshold, $E_{IS}[L; L > b]$. The proposed algorithm is applied to the set of portfolios in Glasserman

EXHIBIT 17
Results for $\hat{\gamma}_{IS}$ of Portfolio (b.1)

Loss	$\hat{\gamma}$	$\hat{\gamma}_{IS}$	$SE(\hat{\gamma})$	$SE(\hat{\gamma}_{IS})$	$CV(\hat{\gamma})$	$CV(\hat{\gamma}_{IS})$	*VRR*
60	$4.008*10^{-1}$	$3.833*10^{-2}$	$1.366*10^{-2}$	$4.810*10^{-3}$	$3.408*10$	$8.873*10^{-1}$	$1.612*10^{3}$
90	$6.295*10^{-2}$	$5.924*10^{-2}$	$1.393*10^{-2}$	$7.420*10^{-4}$	$2.213*10^{2}$	$8.855*10^{-1}$	$7.051*10^{4}$
120	$3.387*10^{-2}$	$3.110*10^{-2}$	$1.337*10^{-2}$	$4.016*10^{-4}$	$3.946*10^{2}$	$9.130*10^{-1}$	$2.215*10^{5}$
150	$5.781*10^{-3}$	$4.367*10^{-3}$	$9.604*10^{-3}$	$6.316*10^{-5}$	$1.661*10^{3}$	1.023	$4.625*10^{6}$
180	$6.183*10^{-4}$	$5.783*10^{-4}$	$2.820*10^{3}$	$9.842*10^{-6}$	$4.560*10^{3}$	1.203	$1.641*10^{7}$

EXHIBIT 18
Results for $\hat{\gamma}_{IS}$ of Portfolio (b.2)

Loss	$\hat{\gamma}$	$\hat{\gamma}_{IS}$	$SE(\hat{\gamma})$	$SE(\hat{\gamma}_{IS})$	$CV(\hat{\gamma})$	$CV(\hat{\gamma}_{IS})$	*VRR*
70	$2.109*10^{-1}$	$2.041*10^{-1}$	$1.605*10^{-2}$	$1.875*10^{-3}$	$7.612*10$	$6.497*10^{-1}$	$1.465*10^{4}$
80	$1.267*10^{-1}$	$1.183*10^{-1}$	$1.614*10^{-2}$	$1.191*10^{-3}$	$1.274*10^{2}$	$7.117*10^{-1}$	$3.675*10^{4}$
90	$7.414*10^{-2}$	$6.799*10^{-2}$	$1.625*10^{-2}$	$7.320*10^{-4}$	$2.192*10^{2}$	$7.612*10^{-1}$	$9.855*10^{4}$
100	$4.427*10^{-2}$	$3.873*10^{-2}$	$1.641*10^{-2}$	$4.431*10^{-4}$	$3.707*10^{2}$	$8.090*10^{-1}$	$2.742*10^{5}$
110	$2.554*10^{-2}$	$2.214*10^{-1}$	$1.633*10^{-2}$	$2.882*10^{-4}$	$6.396*10^{2}$	$9.207*10^{-1}$	$6.421*10^{5}$

EXHIBIT 19

Results for $\hat{\gamma}_{IS}$ of Portfolio (b.3)

Loss	$\hat{\gamma}$	$\hat{\gamma}_{IS}$	$SE(\hat{\gamma})$	$SE(\hat{\gamma}_{IS})$	$CV(\hat{\gamma})$	$CV(\hat{\gamma}_{IS})$	*VRR*
120	$2.1*10^{-1}$	$2.193*10^{-1}$	$2.389*10^{-2}$	$5.655*10^{-4}$	7.476	$3.801*10^{-1}$	$3.569*10^{5}$
140	$1.058*10^{-1}$	$1.077*10^{-1}$	$2.356*10^{-2}$	$2.573*10^{-4}$	$2.402*10$	$2.458*10^{-1}$	$1.676*10^{6}$
160	$5.036*10^{-2}$	$5.212*10^{-2}$	$2.295*10^{-2}$	$1.194*10^{-4}$	$1.137*10^{2}$	$1.823*10^{-1}$	$7.395*10^{6}$
180	$2.490*10^{-2}$	$2.499*10^{-2}$	$2.103*10^{-2}$	$5.750*10^{-5}$	$2.226*10^{2}$	$1.690*10^{-1}$	$2.676*10^{7}$
200	$1.223*10^{-2}$	$1.185*10^{-2}$	$1.824*10^{-2}$	$2.881*10^{-5}$	$4.557*10^{2}$	$1.619*10^{-1}$	$8.016*10^{7}$

et al. [2000] with a variety of characteristics to validate its efficiency. The numerical results demonstrate the importance sampling estimator is unbiased compared to the naive Monte Carlo method. Improvements of more than two orders of magnitude and often more than three or four orders of magnitude were obtained. The numerical results also exhibit that the algorithm guarantees variance reduction, which increases as the loss threshold b is moved to extreme values. Dramatic variance reduction is obtained for portfolio with high correlation between the risk factors. The proposed algorithm has a constant coefficient of variation, suggesting the importance sampling estimator has bounded relative error.

The central idea of our algorithm is to ensure the loss event ($L > b$) always occurs in every sample path. Lemmas 2, 3, and 4 establish the guidelines for choosing an appropriate importance sampling distribution wisely. Our algorithm allows more general problem setting for the risk factor distribution (Hull and White [1998]), as long as the dependency structure of the risk factors is driven by a multivariate normal distribution. Proposition 5 proves that our importance sampling estimator definitely decreases variance compared with naive Monte Carlo simulation.

Our algorithm appears to be straightforward and allows the practical needs of risk management to be realized easily. We can easily observe the required parameters for valuation and simulation in the front office of a bank. The nonlinear characteristic of the option portfolios makes the problems of interest more complex than linear portfolios. Moreover, the Basel Committee on Banking Supervision [2011] has pointed out that the extreme losses occurring from rare events in a stressed market should be highly valued, particularly when dealing with derivatives. The delta-gamma method is an alternative solution for speeding up, but the drawback is that the approximation provides only partial valuation in simulation. Conclusively, this article establishes a fast Monte Carlo algorithm using importance sampling technique for estimating VaR and ES that is suitable under the full valuation framework.

ENDNOTES

The authors thank the Risk and Insurance Research Center of National Chengchi University for its financial support. Ming-Hua Hsieh is grateful to the Ministry of Science and Technology of Taiwan for its partial financial support under contract No. 99-2410-H-004-086.

[1]The stochastic differential equation provides a straightforward way of establishing the risk factor model. See Hull [2011] for more discussions.

[2]The term "pdf" means probability distribution function; "cdf" is cumulative distribution function.

REFERENCES

Alexander, C. *Market Risk Analysis, Value-at-Risk Models*. Wiley, 2009.

Artzner, P., F. Delbaen, J.M. Eber, and D. Heath. "Think Coherently." *Risk*, 10 (1997), pp. 68-71.

——. "Coherent Measure of Risk." *Mathematical Finance*, 9 (1999), pp. 203-228.

Asmussen, S., and P. Glynn. *Stochastic Simulation: Algorithms and Analysis*. Springer, 2007.

Basel Committee on Banking Supervision. "Revisions to the Basel II Market Risk Framework." 2011. Available at http://www.bis.org.

Chen, Z., and P. Glasserman. "Fast Pricing of Basket Default Swaps." *Operations Research*, 56 (2008), pp. 286-303.

Chiang, M.-H., M.-L. Yueh, and M.-H. Hsieh, "An Efficient Algorithm for Basket Default Swap Valuation." *The Journal of Derivatives*, Vol. 15, No. 2 (2007), pp. 8-19.

Glasserman, P. *Monte Carlo Methods in Financial Engineering.* New York: Springer Verlag, 2004.

Glasserman, P., P. Heidelberger, and P. Shahabuddin. "Variance Reduction Techniques for Estimating Value-at-Risk." *Management Science*, 46 (2000), pp. 1349-1364.

——. "Portfolio Value-At-Risk with Heavy-Tailed Risk Factors." *Mathematical Finance*, 12 (2002), pp. 239-269.

Glynn, P., and D.L. Iglehart. "Importance Sampling for Stochastic Simulations." *Management Science*, 35 (1989), pp. 1367-1392.

Hammersley, J., and D. Handscomb. *Monte Carlo Methods.* 1964.

Heidelberger, P. "Fast Simulation of Rare Events in Queuing and Reliability Models." *ACM Transactions on Modeling and Computer Simulation*, Vol. 5, No. 1 (1995), pp. 43-85.

Hoogerheide, L., and H.K. van Dijk. "Bayesian Forecasting of Value at Risk and Expected Shortfall Using Adaptive Importance Sampling." *International Journal of Forecasting*, Vol. 26, No. 2 (2010), pp. 231-247.

Hull, J. *Options, Futures and Other Derivatives*, 8th edition. Prentice Hall, 2011.

——. *Risk Management and Financial Institutions*, 3rd edition. Wiley Finance, 2012.

Hull, J., and A. White. "Value at Risk when Daily Changes Are Not Normally Distributed." *The Journal of Derivatives*, 5 (1998), pp. 9-19.

Jorion, P. *Value at Risk: The New Benchmark for Managing Financial Risk*, 3rd edition. New York: McGraw-Hill, 2007.

Joshi, M., and D. Kainth. "Rapid and Accurate Development of Prices and Greeks for nth to Default Credit Swaps in the Li Model." *Quantitative Finance*, Vol. 4, No. 3 (2004), pp. 266-275.

Longerstaey, J. *RiskMetrics™—Technical Document*, 4th edition. J.P. Morgan/Reuters, 1996.

Rouvinez, C. "Going Greek with VAR." *Risk*, Vol. 10, No. 2 (1997), pp. 57-65.

Wilson, T. *Value-at-Risk*. Chichester, U.K.: Wiley, 1999.

Yueh, M.-L., Wong, M.C.W. "Analytical VaR and Expected Shortfall for Quadratic Portfolios." *The Journal of Derivatives*, Vol. 2010, No. 1 (2010), pp. 1-12.