Kernel Estimation of the Greeks for Options with Discontinuous Payoffs

Guangwu Liu
Department of Management Sciences, City University of Hong Kong, Kowloon, Hong Kong, China,
msgw.liu@cityu.edu.hk

L. Jeff Hong
Department of Industrial Engineering and Logistics Management, The Hong Kong University of Science and Technology,
Clear Water Bay, Hong Kong, China, hongl@ust.hk

The Greeks are the derivatives (also known as sensitivities) of the option prices with respect to market parameters. They play an important role in financial risk management. Among many Monte Carlo methods of estimating the Greeks, the classical pathwise method requires only the pathwise information that is directly observable from simulation and is generally easier to implement than many other methods. However, the classical pathwise method is generally not applicable to the Greeks of options with discontinuous payoffs and the second-order Greeks. In this paper, we generalize the classical pathwise method to allow discontinuity in the payoffs. We show how to apply the new pathwise method to the first- and second-order Greeks and propose kernel estimators that require little analytical efforts and are very easy to implement. The numerical results show that our estimators work well for practical problems.

Subject classifications: finance; securities.
Area of review: Financial Engineering.
History: Received August 2008; revisions received August 2009, November 2009; accepted February 2010. Published online in *Articles in Advance* December 7, 2010.

1. Introduction

In financial risk management, the Greeks are the derivatives (also known as sensitivities) of the option prices with respect to market parameters. For instance, the first-order derivatives with respect to the price and volatility of the underlying are called `delta` and `vega`, respectively, and the second-order derivative with respect to the price of the underlying is called `gamma`. The Greeks play an important role in financial risk management. Each of the Greeks measures a different dimension of the market risk in an option position, and traders manage the Greeks so that all risks are acceptable (Hull 2006). They are important in hedging. For instance, `delta` indicates the number of units of the underlying security to hold in the hedged portfolio, and `gamma` is used to determine the optimal time interval required for rebalancing a hedge under transaction costs (Broadie and Glasserman 1996). As pointed out by Glasserman (2004, p. 377), "whereas the [option] prices themselves can often be observed in the market, their sensitivities cannot, so accurate calculation of sensitivities is arguably even more important than calculation of prices." In this paper we consider how to calculate the Greeks using Monte Carlo simulation.

The price of an option can be written as an expectation of the discounted payoff of the option, where the price

dynamics of the underlying is modeled as a stochastic process under the risk-neutral measure (Chen and Hong 2007). Then the Greeks are the derivatives of the expectation with respect to the parameters of the stochastic process. In the literature several derivative-estimation methods can be applied to estimate the Greeks, including finite-difference approximations, the classical pathwise method,[1] the likelihood ratio (LR) method, the weak derivative (WD) method, and the Malliavin calculus method. For comprehensive reviews, readers are referred to Glasserman (2004), Fu (2006), and Asmussen and Glynn (2007).

Finite-difference approximations are easy to implement, but they require simulating at multiple values, and they typically lead to estimators that have large mean-squared errors (Glasserman 2004). The classical pathwise method (see, e.g., Ho and Cao 1983, and Broadie and Glasserman 1996) often has the best performance among all methods when it is applicable, whereas its major limitation lies in that it is not applicable to options with discontinuous payoffs and the second-order Greeks. The LR method (see, e.g., Glynn 1987, and Broadie and Glasserman 1996) can handle options with discontinuous payoffs, although its drawbacks include that it often has a large variance and that it requires knowing the densities of the underlying price dynamics. The WD method (see Pflug 1988, and Pflug and Weisshaupt 2005) can also handle options with discontinuous payoffs,

although a significant drawback is its requirement of many resimulations, which limits its practical use. The Malliavin calculus method (see, e.g., Bernis et al. 2003, and Chen and Glasserman 2007) is essentially a combination of the classical pathwise method and the LR method, although its popularity is limited by its requirement of the heavy machinery of Malliavin calculus.

In general, the classical pathwise method is easy to implement and has a small variance, and hence is more preferable than other methods. However, it is not applicable to options with discontinuous payoffs, which significantly limits its applicability. To overcome this difficulty, Fu and Hu (1997) suggested conditioning on certain random variables to smooth the discontinuous payoffs, which is called smoothed perturbation analysis (SPA). Early work related to SPA includes Gong and Ho (1987), and readers are referred to Asmussen and Glynn (2007) for a recent overview. However, SPA requires finding the appropriate random variables on which to condition, and it may be difficult to implement in practice.

In this paper, we circumvent the difficulty of discontinuous payoffs from another aspect. By extending the results in Hong (2009) and Hong and Liu (2010), we first convert the Greek into the sum of an ordinary expectation and a derivative with respect to an auxiliary parameter. We then estimate the expectation and the derivative by a sample mean and a kernel estimator, respectively. The kernel method has been studied extensively in the literature of nonparametric regression, and readers are referred to Bosq (1998) for a comprehensive overview. Recently, the kernel method has also been used by Elie et al. (2007) to study Greek estimation on a different focus. Whereas we employ the kernel method to extend the classical pathwise method to options with discontinuous payoffs and the second-order Greeks, they employ the kernel method to extend the LR method to cases where densities of the price dynamics are not explicitly known. To do so, they first randomize the parameter by introducing a priori distribution. They then use the kernel method to estimate score functions of the LR method and devise a simpler estimator by integration by parts argument.

To summarize, our contributions are twofold. First, we generalize the classical pathwise method to allow discontinuity in the payoffs. Second, we devise kernel estimators for both first- and second-order Greeks. Implementation of the proposed estimators is very easy and requires only the pathwise derivatives, which are usually readily computable from simulation. We run numerical tests on two options, including an Asian digital option and a barrier option. We further apply our approach to estimate price sensitivities of portfolio credit derivatives. The examples show that the proposed estimators have good performances.

The rest of the paper is organized as follows: The kernel estimation, discussions on technical details, and examples are discussed in §§2 and 3 for the first- and second-order Greeks, respectively. We report the numerical results in §4.

Extension of our approach to sensitivity analysis of portfolio credit derivatives is provided in §5, followed by the conclusions in §6. Some lengthy proofs and discussions are in the appendix and the electronic companion. The electronic companion is part of the online version that can be found at http://or.journal.informs.org/.

2. The First-Order Greeks

Let S_t denote the price dynamics of the underlyings at time $t \geq 0$ under the risk-neutral measure. Let $\{0 = t_0 < t_1 < \cdots < t_k = T\}$ be the time points between time 0 and time T. When we simulate S_t, we often discretize the process and simulate it at t_i, $i = 1, \ldots, k$ (Glasserman 2004). For simplicity of the notation, we let S_i denote S_{t_i} for all $i = 1, 2, \ldots, k$, and let $S = (S_1, \ldots, S_k)^T$.

Throughout the paper we are interested in an option whose discounted payoff function can be written as $f(S) = g(S) \cdot 1_{\{h(S) \geq 0\}}$, where g and h are differentiable functions. Here $g(S)$ is the discounted cash flow of the option if it is exercised, and $h(S) \geq 0$ defines the exercising condition. Sometimes, the exercising condition may be complicated and involve several individual conditions, i.e., $f(S) = g(S) \cdot 1_{\{h_1(S) \geq 0\}} \cdots 1_{\{h_m(S) \geq 0\}}$. If we define $h(S) = \min\{h_1(S), \ldots, h_m(S)\}$, then we can still write $f(S)$ in the form of $f(S) = g(S) \cdot 1_{\{h(S) \geq 0\}}$. For instance, the discounted payoff function of an up-and-out barrier call option can be written as $f(S) = e^{-rT}(S_k - K) \cdot 1_{\{S_k \geq K\}} \cdot 1_{\{\max\{S_1, \ldots, S_k\} \leq U\}}$, where r is the risk-free borrowing rate, K is the strike price and U is the barrier. Then, we may write $g(S) = e^{-rT}(S_k - K)$ and $h(S) = \min\{S_k - K, U - \max\{S_1, \ldots, S_k\}\}$. Let $p = E[f(S)]$, where the expectation is taken with respect to the risk-neutral measure. Then, p is the price of the option.

Note that S may depend on some market parameters, e.g., the initial price S_0, price volatility σ, and risk-free borrowing rate r. Without loss of generality, we let θ denote the parameter in which we are interested and assume that θ is one dimensional and $\theta \in \Theta$ where Θ is an open set. If θ is multidimensional, we may treat each dimension as a one-dimensional parameter while fixing other dimensions constants. Because S is a random vector that depends on θ, $p = E[f(S)]$ is also a function of θ. We denote it as $p(\theta)$. Then, $p'(\theta) = dp(\theta)/d\theta$ is the Greek with respect to θ. For instance, $p'(\theta)$ is known as delta if $\theta = S_0$, vega if $\theta = \sigma$, and rho if $\theta = r$. In this section, we show how to develop a general scheme to estimate $p'(\theta)$. In the later exposition of the paper, for notational simplicity we suppress the dependence of S on θ when there is no confusion.

2.1. Kernel Estimation

To derive kernel estimators for first-order Greeks, we need to use a result that is summarized in the following theorem, whose proof is provided in the appendix. We defer the discussions of the theorem and its assumptions until the next subsection.

THEOREM 1. *Suppose that* $\mathrm{E}[|g(S)|^2] < +\infty$ *and* $\mathrm{E}[|h(S)|^2] < +\infty$. *If the following assumptions are satisfied,*

ASSUMPTION 1. *For any* $\theta \in \Theta$, $g(S)$ *and* $h(S)$ *are differentiable with respect to* θ *with probability 1 (w.p.1), and there exist random variables* K_g *and* K_h *with finite second moments that may depend on* θ, *such that* $|g(S(\theta + \Delta\theta)) - g(S(\theta))| \leqslant K_g|\Delta\theta|$ *and* $|h(S(\theta + \Delta\theta)) - h(S(\theta))| \leqslant K_h|\Delta\theta|$ *when* $|\Delta\theta|$ *is small enough.*[2]

ASSUMPTION 2. *For any* $\theta \in \Theta$, $\partial_\theta \phi(\theta, y)$ *exists and is continuous at* $(\theta, 0)$, *where* $\phi(\theta, y) = \mathrm{E}[g(S) \cdot 1_{\{h(S) \geqslant y\}}]$,

then

$$
\begin{aligned}
p'(\theta) = {} & \mathrm{E}[\partial_\theta g(S) \cdot 1_{\{h(S) \geqslant 0\}}] \\
& - \partial_y \mathrm{E}[g(S)\partial_\theta h(S) \cdot 1_{\{h(S) \geqslant y\}}]|_{y=0}.
\end{aligned} \tag{1}
$$

The critical value of Equation (1) is that it converts $p'(\theta)$, which is a derivative with respect to θ, to a sum of an ordinary expectation and a derivative with respect to y. Note that the derivative with respect to y is much easier to handle than the derivative with respect to θ, because y is not a part of the simulation that generates $g(S)$ and $h(S)$. Therefore, we can use the finite-difference approximations to estimate the derivative without running simulations at multiple input values, which is one of the most significant drawbacks of the typical finite-difference approximations.

Specifically, the second term on the right-hand side of Equation (1) can be estimated by the finite-difference method, which is based on the following relationship:

$$
\begin{aligned}
& -\partial_y \mathrm{E}[g(S)\partial_\theta h(S) \cdot 1_{\{h(S) \geqslant y\}}]|_{y=0} \\
& = -\lim_{\delta \to 0} \frac{1}{\delta} \{ \mathrm{E}[g(S)\partial_\theta h(S) \cdot 1_{\{h(S) \geqslant \delta/2\}}] \\
& \qquad\qquad - \mathrm{E}[g(S)\partial_\theta h(S) \cdot 1_{\{h(S) \geqslant -\delta/2\}}] \} \\
& = \lim_{\delta \to 0} \frac{1}{\delta} \mathrm{E}[g(S)\partial_\theta h(S) \cdot 1_{\{-\delta/2 \leqslant h(S) \leqslant \delta/2\}}].
\end{aligned} \tag{2}
$$

This motivates the use of the kernel method, which is a generalization of the finite-difference method. A kernel Z is a symmetric density such that $uZ(u) \to 0$ as $|u| \to \infty$ and $\int_{-\infty}^{\infty} u^2 Z(u)\, du < \infty$ (Bosq 1998). For instance, the standard normal density function is a widely used kernel. Then by Bosq (1998), we have

$$
\begin{aligned}
& -\partial_y \mathrm{E}[g(S)\partial_\theta h(S) \cdot 1_{\{h(S) \geqslant y\}}]|_{y=0} \\
& = \lim_{\delta \to 0} \frac{1}{\delta} \mathrm{E}\left[g(S)\partial_\theta h(S) \cdot Z\left(\frac{h(S)}{\delta}\right) \right].
\end{aligned} \tag{3}
$$

When $Z(u) = 1_{\{-1/2 \leqslant u \leqslant 1/2\}}$, which is the density of a uniform $(-1/2, 1/2)$ distribution and is known as a uniform kernel, the right-hand sides of Equations (2) and (3) are the same. Therefore, the finite-difference method we developed is a special case of the kernel method with a uniform kernel. In this paper, we suggest using a smooth kernel, e.g.,

the standard normal density, instead of a uniform kernel, because the estimator is more robust with a smooth kernel (Bosq 1998).

There is another approach[3] to explain Theorem 1 and Equation (3). Let $F(u) = \int_{-\infty}^{u} Z(v)\, dv$. Then, $F(u)$ is the cumulative distribution function of the random variable whose density is the kernel function $Z(u)$. Note that $1_{\{x \geqslant 0\}} = \lim_{\delta \to 0} F(x/\delta)$ almost everywhere when $\delta > 0$. Then,

$$
p(\theta) = \mathrm{E}[g(S) \cdot 1_{\{h(S) \geqslant 0\}}] = \lim_{\delta \to 0} \mathrm{E}\left[g(S) \cdot F\left(\frac{h(S)}{\delta}\right) \right].
$$

Suppose that we can interchange the differentiation and limit (which may be difficult to verify),

$$
\begin{aligned}
p'(\theta) & = \frac{d}{d\theta} \lim_{\delta \to 0} \mathrm{E}\left[g(S) \cdot F\left(\frac{h(S)}{\delta}\right) \right] \\
& = \lim_{\delta \to 0} \frac{d}{d\theta} \mathrm{E}\left[g(S) \cdot F\left(\frac{h(S)}{\delta}\right) \right] \\
& = \lim_{\delta \to 0} \mathrm{E}\left[\partial_\theta g(S) \cdot F\left(\frac{h(S)}{\delta}\right) \right] \\
& \quad + \lim_{\delta \to 0} \frac{1}{\delta} \mathrm{E}\left[g(S)\partial_\theta h(S) \cdot Z\left(\frac{h(S)}{\delta}\right) \right] \\
& = \mathrm{E}[\partial_\theta g(S) \cdot 1_{\{h(S) \geqslant 0\}}] \\
& \quad + \lim_{\delta \to 0} \frac{1}{\delta} \mathrm{E}\left[g(S)\partial_\theta h(S) \cdot Z\left(\frac{h(S)}{\delta}\right) \right].
\end{aligned} \tag{4}
$$

Note that the second term on the right-hand side of Equation (4) is the same as the right-hand side of Equation (3). Then, by Equation (3), we arrive at the same conclusion as in Theorem 1.

Let (g_l, h_l, g'_l, h'_l) denote the lth independent observation of $(g(S), h(S), \partial_\theta g(S), \partial_\theta h(S))$, $l = 1, 2, \ldots, n$, obtained from a simulation. Based on Equation (1), we propose the following estimator of $p'(\theta)$,

$$
\overline{G}_n = \frac{1}{n} \sum_{l=1}^{n} g'_l \cdot 1_{\{h_l \geqslant 0\}} + \frac{1}{n\delta_n} \sum_{l=1}^{n} g_l \cdot h'_l \cdot Z\left(\frac{h_l}{\delta_n}\right), \tag{5}
$$

where $\delta_n > 0$ is called a bandwidth parameter in the kernel estimation.

Our kernel estimation method can be viewed as a generalization of the classical pathwise method, because it uses only the observations of $(g(S), h(S), \partial_\theta g(S), \partial_\theta h(S))$ that can be obtained directly from the simulated sample paths. It is unlike the likelihood ratio method or the weak derivative method, where the explicit form of the density needs to be used to compute the estimators. Therefore, it is as simple to implement in practice as the classical pathwise method.

Essentially, the proposed kernel method is a combination of the classical pathwise method and the kernel-smoothing technique. From this perspective, the proposed method is closely related to SPA that uses conditional

Monte Carlo as a smoothing technique rather than kernel smoothing. Although kernel smoothing produces biased estimators whereas conditional Monte Carlo produces unbiased ones, kernel smoothing has an attractive advantage that it is far easier to implement than conditional Monte Carlo. Furthermore, as pointed out in the later discussions, both Assumptions 1 and 2 are typically satisfied in practice. Therefore, $\overline{G}_n$ is much more broadly applicable than classical pathwise estimators and SPA estimators.

The asymptotic properties of $\overline{G}_n$ are standard in the literature of the kernel method (e.g., Bosq 1998). For completeness of the paper, we prove in the electronic companion that $\overline{G}_n$ is a consistent estimator and follows a central limit theorem if $\delta_n \to 0$ and $n\delta_n \to \infty$ as $n \to \infty$, and the optimal rate of convergence of the estimator is $n^{-2/5}$ and it is achieved when $\delta_n = cn^{-1/5}$ for some constant $c > 0$. Furthermore, in the electronic companion, we also provide a procedure on how to select δ_n based on a pilot simulation run.

2.2. Discussions

This subsection is devoted to the discussions of Assumptions 1 and 2, and the connections between Theorem 1 and some other results in the literature.

Assumption 1 in Theorem 1 is a typical assumption in the pathwise sensitivity estimation. Note that when S follows commonly used models of price dynamics, $S(\theta)$ is continuously differentiable and $\|S(\theta + \Delta\theta) - S(\theta)\| \leqslant K_S|\Delta\theta|$ for some random variable K_S (Broadie and Glasserman 1996), where $\|\cdot\|$ denotes the Euclidean norm. Therefore, to satisfy Assumption 1, we need that g and h are Lipschitz continuous and differentiable almost everywhere (a.e.). For most financial options, if not all of them, g and h are Lipschitz continuous and differentiable a.e. Even when $h(\cdot) = \min\{h_1(\cdot), \ldots, h_m(\cdot)\}$, h is Lipschitz continuous and differentiable a.e., if $h_1, \ldots, h_m$ are Lipschitz continuous and differentiable a.e. Therefore, *Assumption 1 is typically satisfied in practice.*

Assumption 2 is critical to the development of our method. Let $f_{S(\theta)}(s)$ denote the density of $S(\theta)$. We propose the following lemma to verify the assumption. The proof of the lemma is provided in the appendix.

LEMMA 1. *Suppose that, for any $\theta \in \Theta$, $f_{S(\theta)}(s)$ is continuously differentiable in θ for almost all $s \in \Re^k$, and there exists a function $\beta(s)$ such that $|f_{S(\theta + \Delta\theta)}(s) - f_{S(\theta)}(s)| \leqslant \beta(s)|\Delta\theta|$ when $|\Delta\theta|$ is small enough, and $\int_{\Re^k} |g(s)| \beta(s)\,ds < +\infty$. Then, Assumption 2 is satisfied.*

Note that the conditions of Lemma 1 are typical conditions for the likelihood ratio method to be applicable to estimating $d\mathrm{E}[g(S)]/d\theta$ or $p'(\theta)$ (see, e.g., Asmussen and Glynn 2007 and L'Ecuyer 1991). As pointed out by Broadie and Glasserman (1996), these conditions are typically satisfied in the context of estimating the Greeks. Therefore, Lemma 1 shows that *Assumption 2 is typically satisfied in practice.*

Theorem 1 generalizes several results in the literature. For instance, we may let $h(S) = R$ where R is a uniform $(0, 1)$ random number. Then, $h(S) \geqslant 0$ w.p.1 and $\partial_\theta h(S) = 0$. By Equation (1),

$$p'(\theta) = \frac{d}{d\theta}\mathrm{E}[g(S)] = \mathrm{E}[\partial_\theta g(S)],$$

which is the conclusion of the classical pathwise method of Broadie and Glasserman (1996). If we let $g(S) = 1$, then $\partial_\theta g(S) = 0$, and by Equation (1),

$$p'(\theta) = \frac{d}{d\theta}\Pr\{h(S) \geqslant 0\} = -\partial_y \mathrm{E}[\partial_\theta h(S) \cdot 1_{\{h(S) \geqslant y\}}]|_{y=0},$$

which is a result of the probability sensitivity derived in Hong and Liu (2010).

Furthermore, if we take one step further from Theorem 1, we may express the first-order Greek $p'(\theta)$ in terms of densities and conditional expectations. To be specific, we let f_h denote the density of $h(S)$. If $\gamma(y) \equiv f_h(y) \cdot \mathrm{E}[g(S)\partial_\theta h(S)|h(S) = y]$ is continuous at $y = 0$, then

$$p'(\theta) = \mathrm{E}[\partial_\theta g(S) \cdot 1_{\{h(S) \geqslant 0\}}]$$
$$+ f_h(0) \cdot \mathrm{E}[g(S)\partial_\theta h(S)|h(S) = 0]. \quad (6)$$

Here the requirement that $\gamma(y)$ is continuous at $y = 0$ will not impose any practical obstacle on the result. Note that both the density and conditional expectation are defined by integrations (Durrett 2005). Then, mathematically $\gamma(y)$ can take any arbitrary value at $y = 0$ because $\{h(S) = 0\}$ is a probability-zero event. Therefore, we need to assume that it is continuous in order to prove Equation (6). By Assumption 2, we can see that $\partial_y \mathrm{E}[g(S)\partial_\theta h(S) \cdot 1_{\{h(S) \geqslant y\}}]$ is continuous at $y = 0$, because from the way of showing Equation (1) we actually have

$$\partial_y \mathrm{E}[g(S)\partial_\theta h(S) \cdot 1_{\{h(S) \geqslant y\}}]$$
$$= \partial_\theta \phi(\theta, y) - \mathrm{E}[\partial_\theta g(S) \cdot 1_{\{h(S) \geqslant y\}}],$$

where both terms in the right-hand side of the above equation are continuous at $y = 0$. We may define $\tilde{\gamma}(y) = \partial_y \mathrm{E}[g(S)\partial_\theta h(S) \cdot 1_{\{h(S) \geqslant y\}}]$ in a neighborhood of $y = 0$. Then, $\tilde{\gamma}(y)$ is continuous and $\tilde{\gamma}(y) = \gamma(y)$ a.e. in the neighborhood of $y = 0$. For practically used definitions of the density and conditional expectation, $\gamma(y)$ is typically the same as $\tilde{\gamma}(y)$. Then, $\gamma(y)$ is continuous at $y = 0$, and hence Equation (6) holds.

2.3. Examples

We use two examples to illustrate how to apply Theorem 1 and the kernel estimator $\overline{G}_n$ to estimate the first-order Greeks of options with discontinuous payoffs.

2.3.1. Asian Digital Option. The payoff function of an Asian digital option can be written as $f(S) = e^{-rT} \cdot 1_{\{\bar{S} \geqslant K\}}$, where $\bar{S} = \frac{1}{k}\sum_{i=1}^{k} S_i$. Then, we let $g(S) = e^{-rT}$ and $h(S) = \bar{S} - K$. It is easy to verify Assumptions 1 and 2 if S follows a commonly used model of price dynamics, e.g., a geometric Brownian motion. Suppose that we are interested in estimating delta, i.e., dp/dS_0. By Theorem 1,

$$\frac{dp}{dS_0} = -\partial_y E\left[e^{-rT}\frac{d\bar{S}}{dS_0} \cdot 1_{\{\bar{S}-K \geqslant y\}} \right]\Big|_{y=0}.$$

We apply the estimator $\bar{G}_n$ of Equation (5) to estimate delta:

$$\bar{G}_n = \frac{1}{n\delta_n}\sum_{l=1}^{n} e^{-rT}\frac{d\bar{S}_l}{dS_0} Z\left(\frac{\bar{S}_l - K}{\delta_n} \right),$$

where $(\bar{S}_l, d\bar{S}_l/dS_0)$ denotes the lth observation of $(\bar{S}, d\bar{S}/dS_0)$.

2.3.2. Up-and-Out Barrier Call Option. We consider an up-and-out barrier call option whose discounted payoff function can be written as

$$f(S) = e^{-rT}(S_k - K) \cdot 1_{\{S_k \geqslant K\}} \cdot 1_{\{\check{S} \leqslant U\}}$$
$$= e^{-rT}(S_k - K) \cdot 1_{\{\min\{S_k-K, U-\check{S}\} \geqslant 0\}},$$

where $\check{S} = \max\{S_1, \ldots, S_k\}$. Then, we define $g(S) = e^{-rT}(S_k - K)$ and $h(S) = \min\{S_k - K, U - \check{S}\}$. If S is Lipschitz continuous in θ, then $\check{S}$ is also Lipschitz continuous in θ. Therefore, Assumptions 1 and 2 can be easily verified if S follows a commonly used model of price dynamics, e.g., a geometric Brownian motion.

Suppose that we are interested in estimating delta, i.e., dp/dS_0. Then by Theorem 1, we can derive that

$$\frac{dp}{dS_0} = E\left[e^{-rT}\frac{dS_k}{dS_0}1_{\{S_k \geqslant K\}}1_{\{\check{S} \leqslant U\}} \right]$$
$$+ \partial_y E\left[e^{-rT}(S_k - K)\frac{d\check{S}}{dS_0}1_{\{S_k - K \geqslant 0\}}1_{\{U-\check{S} \geqslant y\}} \right]\Big|_{y=0}. \quad (7)$$

For details of the derivation, interested readers are referred to §A.4 of the appendix.

We apply the estimator $\bar{G}_n$ of Equation (5) to estimate delta:

$$\bar{G}_n = \frac{1}{n}\sum_{l=1}^{n} e^{-rT}S'_{k,l} \cdot 1_{\{S_{k,l} \geqslant K\}}1_{\{\check{S}_l \leqslant U\}}$$
$$- \frac{1}{n\delta_n}\sum_{l=1}^{n} e^{-rT}(S_{k,l} - K)\check{S}'_l 1_{\{S_{k,l} \geqslant K\}}Z\left(\frac{U - \check{S}_l}{\delta_n} \right),$$

where $(S_{k,l}, S'_{k,l}, \check{S}_l, \check{S}'_l)$ denotes the lth observation of $(S_k, dS_k/dS_0, \check{S}, d\check{S}/dS_0)$.

3. The Second-Order Greeks

In this section, we consider the second-order Greeks of options with discontinuous payoffs. Because second-order Greeks are essentially second-order derivatives of the option price, they may involve two market parameters θ_1 and θ_2 on which S may depend. We abuse the notation Θ defined in §2 to denote an open set in the $\Re^2$ space, and assume that $(\theta_1, \theta_2) \in \Theta$. In this case, we write $p(\theta_1, \theta_2) = E[f(S)]$ to denote the option price, and we are interested in estimating $\partial_{\theta_2}\partial_{\theta_1}p(\theta_1, \theta_2)$.

3.1. Kernel Estimation

To derive kernel estimators for the second-order Greeks, we need to use a result that is summarized in the following theorem, whose proof is provided in the appendix. We defer the discussions of the theorem and its assumptions until the next subsection.

THEOREM 2. *Suppose that* $E[|g(S)|^4] < +\infty$ *and* $E[|h(S)|^4] < +\infty$. *If the following assumptions are satisfied,*

ASSUMPTION 3. *For any* $(\theta_1, \theta_2) \in \Theta$, $g(S)$ *and* $h(S)$ *are differentiable with respect to* θ_1 *w.p.1,* $\partial_{\theta_1}g(S)$ *and* $\partial_{\theta_1}h(S)$ *are differentiable with respect to* θ_2 *w.p.1, and there exist random variables* K_g, K_h, L_g, *and* L_c *with finite fourth moments, which may depend on* (θ_1, θ_2), *such that*

$$|g(S(\theta_1 + \Delta\theta_1, \theta_2)) - g(S(\theta_1, \theta_2))| \leqslant K_g|\Delta\theta_1|,$$
$$|g(S(\theta_1, \theta_2 + \Delta\theta_2)) - g(S(\theta_1, \theta_2))| \leqslant K_g|\Delta\theta_2|,$$
$$|h(S(\theta_1 + \Delta\theta_1, \theta_2)) - h(S(\theta_1, \theta_2))| \leqslant K_h|\Delta\theta_1|,$$
$$|h(S(\theta_1, \theta_2 + \Delta\theta_2)) - h(S(\theta_1, \theta_2))| \leqslant K_h|\Delta\theta_2|,$$
$$|\partial_{\theta_1}g(S(\theta_1, \theta_2 + \Delta\theta_2)) - \partial_{\theta_1}g(S(\theta_1, \theta_2))| \leqslant L_g|\Delta\theta_2|,$$

and

$$|g(S(\theta_1, \theta_2 + \Delta\theta_2))\partial_{\theta_1}h(S(\theta_1, \theta_2 + \Delta\theta_2))$$
$$- g(S(\theta_1, \theta_2))\partial_{\theta_1}h(S(\theta_1, \theta_2))| \leqslant L_c|\Delta\theta_2|$$

when $|\Delta\theta_1|$ *and* $|\Delta\theta_2|$ *are small enough.*

ASSUMPTION 4. *For any* $\theta \in \Theta$, $\partial_{\theta_2}\partial_{\theta_1}\phi(\theta_1, \theta_2, y)$ *exists and is continuous at* $(\theta_1, \theta_2, 0)$, *where* $\phi(\theta_1, \theta_2, y) = E[g(S) \cdot 1_{\{h(S) \geqslant y\}}]$,

then

$$\partial_{\theta_2}\partial_{\theta_1}p(\theta_1, \theta_2)$$
$$= E[\partial_{\theta_2}\partial_{\theta_1}g(S) \cdot 1_{\{h(S) \geqslant 0\}}] - \partial_y E[(g(S)\partial_{\theta_2}\partial_{\theta_1}h(S)$$
$$+ \partial_{\theta_1}g(S)\partial_{\theta_2}h(S) + \partial_{\theta_2}g(S)\partial_{\theta_1}h(S)) \cdot 1_{\{h(S) \geqslant y\}}]|_{y=0}$$
$$+ \partial_y^2 E[g(S)\partial_{\theta_1}h(S)\partial_{\theta_2}h(S) \cdot 1_{\{h(S) \geqslant y\}}]|_{y=0}. \quad (8)$$

REMARK 1. For commonly used second-order Greek gamma, i.e., $\theta_1 = \theta_2 = S_0$, we may simplify the notation and have

$$
\begin{aligned}
\partial^2_{S_0} p(S_0) \\
= & E[\partial^2_{S_0} g(S) \cdot 1_{\{h(S) \geqslant 0\}}] \\
& - \partial_y E[(g(S) \partial^2_{S_0} h(S) + 2 \partial_{S_0} g(S) \partial_{S_0} h(S)) \cdot 1_{\{h(S) \geqslant y\}}]|_{y=0} \\
& + \partial^2_y E[g(S)(\partial_{S_0} h(S))^2 \cdot 1_{\{h(S) \geqslant y\}}]|_{y=0}.
\end{aligned}
$$

As shown in the previous section, the second term on the right-hand side of Equation (8) can be estimated by the kernel method. Hence, to estimate $\partial_{\theta_2} \partial_{\theta_1} p(\theta_1, \theta_2)$, a major difficulty is how to estimate the third terms on the right-hand side of Equation (8). Note that similar to the previous analysis,

$$
\begin{aligned}
& - \partial_y E[g(S) \partial_{\theta_1} h(S) \partial_{\theta_2} h(S) \cdot 1_{\{h(S) \geqslant y\}}] \\
& = \lim_{\delta \to 0} \frac{1}{\delta} E\left[g(S) \partial_{\theta_1} h(S) \partial_{\theta_2} h(S) \cdot Z\left(\frac{h(S)-y}{\delta}\right)\right],
\end{aligned}
$$

where Z is a smooth kernel.

Intuitively, by differentiating in y on both sides of the above equation and then letting $y = 0$, we have

$$
\begin{aligned}
& - \partial^2_y E[g(S) \partial_{\theta_1} h(S) \partial_{\theta_2} h(S) \cdot 1_{\{h(S) \geqslant y\}}]|_{y=0} \\
& = - \lim_{\delta \to 0} \frac{1}{\delta^2} E\left[g(S) \partial_{\theta_1} h(S) \partial_{\theta_2} h(S) \cdot Z'\left(\frac{h(S)}{\delta}\right)\right],
\end{aligned}
$$

where $Z'(\cdot)$ denotes the derivative of $Z(\cdot)$.

Therefore, we may propose a kernel estimator for the third term on the right-hand side of Equation (8). Particularly, if we let $(g_l, h_l, \partial_{\theta_1} g_l, \partial_{\theta_2} g_l, \partial_{\theta_2} \partial_{\theta_1} g_l, \partial_{\theta_1} h_l, \partial_{\theta_2} h_l, \partial_{\theta_2} \partial_{\theta_1} h_l)$ denote the lth observation of $(g(S), h(S), \partial_{\theta_1} g(S), \partial_{\theta_2} g(S), \partial_{\theta_2} \partial_{\theta_1} g(S), \partial_{\theta_1} h(S), \partial_{\theta_2} h(S), \partial_{\theta_2} \partial_{\theta_1} h(S))$, then we propose the following estimator of $\partial_{\theta_2} \partial_{\theta_1} p(\theta_1, \theta_2)$:

$$
\begin{aligned}
\bar{H}_n = & \frac{1}{n} \sum_{l=1}^{n} \partial_{\theta_2} \partial_{\theta_1} g_l 1_{\{h_l \geqslant 0\}} \\
& + \frac{1}{n \delta_n} \sum_{l=1}^{n} (g_l \partial_{\theta_2} \partial_{\theta_1} h_l + \partial_{\theta_1} g_l \partial_{\theta_2} h_l + \partial_{\theta_2} g_l \partial_{\theta_1} h_l) Z\left(\frac{h_l}{\delta_n}\right) \\
& + \frac{1}{n \gamma_n^2} \sum_{l=1}^{n} g_l \partial_{\theta_1} h_l \partial_{\theta_2} h_l Z'\left(\frac{h_l}{\gamma_n}\right), \quad (9)
\end{aligned}
$$

where $\delta_n > 0$ and $\gamma_n > 0$ are the bandwidth parameters of the kernel estimation.

The asymptotic properties of $\bar{H}_n$ are standard in the literature of kernel estimation (see, e.g., Bosq 1998). For completeness of the paper, we prove in the electronic companion that $\bar{H}_n$ is a consistent estimator and follows a central limit theorem if $\delta_n \to 0$, $\gamma_n \to 0$, $n \delta_n \to \infty$, and $n \gamma_n^3 \to \infty$ as $n \to \infty$; and the optimal rate of convergence of the estimator is $n^{-2/7}$, and it is achieved when $\delta_n = c_1 n^{-1/5}$ and $\gamma_n = c_2 n^{-1/7}$ for some constants $c_1, c_2 > 0$. Furthermore, in the electronic companion we also provide a procedure on how to select δ_n and γ_n based on a pilot simulation run.

3.2. Discussions

This subsection is devoted to the discussions of the technical details related to Theorem 2.

Assumption 3 in Theorem 2 is a typical assumption in estimating Greeks. It can often be checked. However, generally Assumption 3 may not hold in practice. For instance, when $g(S) = \check{S} = \max\{S_1, \ldots, S_k\}$, there may not exist a random variable L_g with finite fourth moment such that $|\partial_{\theta_1} g(S(\theta_1, \theta_2 + \Delta\theta_2)) - \partial_{\theta_1} g(S(\theta_1, \theta_2))| \leqslant L_g |\Delta\theta_2|$.

REMARK 2. When $g(S)$ and $h(S)$ involve only S_T or $\bar{S}$, Assumption 3 generally holds. However, when $g(S)$ and $h(S)$ involve maxima or minima such as $\check{S} = \max\{S_1, \ldots, S_k\}$ and $\hat{S} = \min\{S_1, \ldots, S_k\}$, Assumption 3 does not hold, except in some special cases, e.g., when $\theta_1 = \theta_2 = S_0$ and S_t follows a linear model of S_0, i.e., $S_t = S_0 \cdot A_t + D_t$, where A_t and D_t are two stochastic processes that are independent of S_0.

Assumption 4 is critical to the analysis of second-order Greeks. Let $f_{S(\theta_1, \theta_2)}(s)$ denote the density function of $S(\theta_1, \theta_2)$. We propose the following lemma to verify the assumption. The proof of the lemma is similar to that of Lemma 1 and is hence omitted.

LEMMA 2. *Suppose that for any $(\theta_1, \theta_2) \in \Theta$, $f_{S(\theta_1, \theta_2)}(s)$ and $\partial_{\theta_1} f_{S(\theta_1, \theta_2)}(s)$ are continuously differentiable in θ_1 and θ_2, respectively, for almost all $s \in \Re^k$, and there exist functions $\beta_1(s)$ and $\beta_2(s)$ such that $|f_{S(\theta_1 + \Delta\theta_1, \theta_2)}(s) - f_{S(\theta_1, \theta_2)}(s)| \leqslant \beta_1(s)|\Delta\theta_1|$ and $|\partial_{\theta_1} f_{S(\theta_1, \theta_2 + \Delta\theta_2)}(s) - \partial_{\theta_1} f_{S(\theta_1, \theta_2)}(s)| \leqslant \beta_2(s)|\Delta\theta_2|$ when $|\Delta\theta_1|$ and $|\Delta\theta_2|$ are small enough, and $\int_{\Re^k} |g(s)|\beta_i(s)\,ds < +\infty$ for $i = 1, 2$. Then, Assumption 4 is satisfied.*

Note that the conditions of Lemma 2 are typical conditions for the likelihood ratio method to be applicable to estimating $\partial_{\theta_2} \partial_{\theta_1} E[g(S)]$ or $\partial_{\theta_2} \partial_{\theta_1} p(\theta_1, \theta_2)$ (see, e.g., Asmussen and Glynn 2007 and L'Ecuyer 1991). As pointed out by Broadie and Glasserman (1996) and Glasserman (2004), these conditions are typically satisfied in the context of estimating the Greeks. Therefore, Lemma 2 shows that *Assumption 4 is typically satisfied in practice.*

Furthermore, if we take one step further from Theorem 2, we may express the second-order Greek $\partial_{\theta_2} \partial_{\theta_1} p(\theta_1, \theta_2)$ in terms of densities and conditional expectations. To be specific, if $f_h(y) \cdot E[g(S) \partial_{\theta_2} \partial_{\theta_1} h(S) + \partial_{\theta_1} g(S) \partial_{\theta_2} h(S) + \partial_{\theta_2} g(S) \partial_{\theta_1} h(S) | h(S) = y]$ and $f_h(y) \cdot E[g(S) \partial_{\theta_1} h(S) \partial_{\theta_2} h(S) | h(S) = y]$ are continuous at $y = 0$, then

$$
\begin{aligned}
& \partial_{\theta_2} \partial_{\theta_1} p(\theta_1, \theta_2) \\
& = E[\partial_{\theta_2} \partial_{\theta_1} g(S) \cdot 1_{\{h(S) \geqslant 0\}}] + f_h(0) \cdot E[g(S) \partial_{\theta_2} \partial_{\theta_1} h(S) \\
& \quad + \partial_{\theta_1} g(S) \partial_{\theta_2} h(S) + \partial_{\theta_2} g(S) \partial_{\theta_1} h(S) | h(S) = 0] \\
& \quad - \partial_y \{f_h(y) \cdot E[g(S) \partial_{\theta_1} h(S) \partial_{\theta_2} h(S) | h(S) = y]\}|_{y=0}. \quad (10)
\end{aligned}
$$

3.3. Examples

We consider again the examples in §2.3 to illustrate how to apply Theorem 2 and the kernel estimator $\bar{H}_n$ to estimate the second-order Greeks of options with discontinuous payoffs.

3.3.1. Asian Digital Option. For the Asian digital option considered in §2.3.1, $g(S) = e^{-rT}$ and $h(S) = \bar{S} - K$. It is easy to verify Assumptions 3 and 4 if S_t follows a commonly used model of price dynamics. Suppose that we are interested in estimating gamma, i.e., d^2p/dS_0^2. By Theorem 2,

$$
\frac{d^2p}{dS_0^2} = -\partial_y \mathrm{E}\left[e^{-rT}\frac{d^2\bar{S}}{dS_0^2}\cdot 1_{\{\bar{S}-K\geqslant y\}}\right]\bigg|_{y=0}
$$
$$
+ \partial_y^2 \mathrm{E}\left[e^{-rT}\left(\frac{d\bar{S}}{dS_0}\right)^2\cdot 1_{\{\bar{S}-K\geqslant y\}}\right]\bigg|_{y=0}.
$$

We apply the estimator $\bar{H}_n$ of Equation (9) to estimate gamma:

$$
\bar{H}_n = \frac{1}{n\delta_n}\sum_{l=1}^{n}e^{-rT}\frac{d^2\bar{S}_l}{dS_0^2}Z\left(\frac{\bar{S}_l-K}{\delta_n}\right)
$$
$$
+ \frac{1}{n\gamma_n^2}\sum_{l=1}^{n}e^{-rT}\left(\frac{d\bar{S}_l}{dS_0}\right)^2 Z'\left(\frac{\bar{S}_l-K}{\gamma_n}\right),
$$

where $(\bar{S}_l, d\bar{S}_l/dS_0, d^2\bar{S}_l/dS_0^2)$ is the lth observation of $(\bar{S}, d\bar{S}/dS_0, d^2\bar{S}/dS_0^2)$.

3.3.2. Up-and-Out Barrier Call Option. For the barrier option considered in §2.3.2, we have $g(S) = e^{-rT}\cdot(S_k-K)$ and $h(S) = \min\{S_k-K, U-\check{S}\}$. Suppose that we are interested in estimating gamma, i.e., $\theta_1 = \theta_2 = S_0$. Note that generally $\partial_{S_0}h(S)$ is not Lipschitz continuous in S_0, and hence Assumption 3 does not hold in general. Therefore, *in general Theorem 2 is not directly applicable to the second-order Greeks of barrier options.*

However, this problem can be resolved when S_t follows a linear model of S_0, e.g., a geometric Brownian motion. By using Theorem 2 we can show that

$$
\frac{d^2p}{dS_0^2} = \mathrm{E}\left[e^{-rT}\frac{d^2S_k}{dS_0^2}1_{\{S_k\geqslant K\}}1_{\{\check{S}\leqslant U\}}\right]
$$
$$
- \partial_y \mathrm{E}\left[e^{-rT}\left(\frac{dS_k}{dS_0}\right)^2 1_{\{U-\check{S}\geqslant 0\}}1_{\{S_k-K\geqslant y\}}\right]\bigg|_{y=0}
$$
$$
+ \partial_y \mathrm{E}\left[e^{-rT}\left(2\frac{dS_k}{dS_0}\frac{d\check{S}}{dS_0}+(S_k-K)\frac{d^2\check{S}}{dS_0^2}\right)\right.
$$
$$
\left.\cdot 1_{\{S_k\geqslant K\}}1_{\{U-\check{S}\geqslant y\}}\right]\bigg|_{y=0}
$$
$$
- \partial_y^2 \mathrm{E}\left[e^{-rT}(S_k-K)\left(\frac{d\check{S}}{dS_0}\right)^2\right.
$$
$$
\left.\cdot 1_{\{S_k\geqslant K\}}1_{\{U-\check{S}\geqslant y\}}\right]\bigg|_{y=0}. \quad (11)
$$

For details of the derivation, interested readers are referred to §A.5 of the appendix.

Then we may use the following kernel estimator to estimate d^2p/dS_0^2:

$$
\bar{H}_n = \frac{1}{n}\sum_{l=1}^{n}e^{-rT}S_{k,l}''1_{\{S_{k,l}\geqslant K\}}1_{\{\check{S}_l\leqslant U\}}
$$
$$
+ \frac{1}{n\delta_n}\sum_{l=1}^{n}e^{-rT}(S_{k,l}')^2 1_{\{\check{S}_l\leqslant U\}}Z\left(\frac{S_{k,l}-K}{\delta_n}\right)
$$
$$
- \frac{1}{n\delta_n}\sum_{l=1}^{n}e^{-rT}[2S_{k,l}\check{S}_l'+(S_{k,l}-K)\check{S}_l'']
$$
$$
\cdot 1_{\{S_{k,l}\geqslant K\}}Z\left(\frac{U-\check{S}_l}{\delta_n}\right)
$$
$$
- \frac{1}{n\gamma_n^2}\sum_{l=1}^{n}e^{-rT}(S_{k,l}-K)(\check{S}_l')^2 1_{\{S_{k,l}\geqslant K\}}Z'\left(\frac{U-\check{S}_l}{\gamma_n}\right),
$$

where $(S_{k,l}, S_{k,l}', S_{k,l}'', \check{S}_l, \check{S}_l', \check{S}_l'')$ denotes the lth observation of $(S_k, dS_k/dS_0, d^2S_k/dS_0^2, \check{S}, d\check{S}/dS_0, d^2\check{S}/dS_0^2)$.

4. Numerical Experiments

To illustrate the performance of the proposed approach ($\bar{G}_n$ in Equation (5) and $\bar{H}_n$ in Equation (9) for first- and second-order Greeks, respectively), we consider the examples in §2.3, including the Asian digital option and the barrier option. For the Asian digital option, existing method including SPA, the LR method and the WD method can also be applied, and numerical results suggest that our approach is significantly better than the existing methods.

For the barrier option, it is not clear how to apply SPA. We show that our approach may outperform the LR and WD methods in many scenarios. In all the examples, the standard normal density is used as the kernel function Z, and δ_n and γ_n are selected by running a pilot simulation[4] according to the procedure as provided in the electronic companion. Results reported are based on 1,000 independent replications.

4.1. Asian Digital Option

Consider an Asian digital option introduced in §2.3.1. Without loss of generality we assume that the discretization points are evenly spaced, and let $\tau = T/k$ denote the time interval. We suppose that S_t follows an Ornstein-Uhlenbeck process, i.e., the price dynamics can be simulated exactly using $S_{i+1} = S_i e^{-b\tau} + \mu(1-e^{-b\tau}) + \sigma\sqrt{(1-e^{-2b\tau})/(2b)}\cdot Z_{i+1}$, for $i = 0, 1, \ldots, k-1$, where b, μ, and σ are mean reversion rate, mean, and volatility, respectively, and $Z_1, \ldots, Z_k$ are independent standard normal random variables.

For this example, analytical price formulas of the Asian digital option and its Greeks can be derived, and hence the exact values of the Greeks can be calculated. In Table 1 we list the exact values of the Greeks for different number

Table 1. Exact values of the Greeks of the Asian digital option.

	delta	vega	theta	gamma
$k=10$	0.9714	0.7411	0.1185	6.0271
$k=20$	1.0609	0.7465	0.1195	6.8791
$k=50$	1.0731	0.7472	0.1196	6.9971

of time steps (k) when $r=5\%$, $\sigma=30\%$, $b=0.2$, $\mu=98$, $S_0=100$, $K=100$, and $T=1$. Then the performance of the proposed method can be evaluated by comparing the proposed estimators to their exact values.

SPA and the LR method can also be applied to this example. An SPA estimator of the first-order Greek can be derived by conditioning on $(S_1,\ldots,S_{k-1})$, i.e.,

$$\frac{\partial}{\partial\theta}\mathrm{E}[1_{\{\bar{S}\geqslant K\}}]=\frac{\partial}{\partial\theta}\mathrm{E}(\mathrm{E}[1_{\{\bar{S}\geqslant K\}}|S_1,\ldots,S_{k-1}])$$

$$=\frac{\partial}{\partial\theta}\mathrm{E}(1-\Phi(Y))=-\mathrm{E}\left[\psi(Y)\cdot\frac{\partial Y}{\partial\theta}\right],$$

where $Y=[kK-(S_1+\cdots+S_{k-2})-(1+e^{-b\tau})S_{k-1}-\mu(1-e^{-b\tau})]/(\sigma\sqrt{(1-e^{-2b\tau})/(2b)})$, and Φ and ψ denote the distribution function and density of the standard normal distribution, respectively. Furthermore, note that $\partial Y/\partial S_0$ is independent of S_0. Then an SPA estimator of gamma can also be derived based on

$$\frac{\partial^2}{\partial S_0^2}\mathrm{E}[1_{\{\bar{S}\geqslant K\}}]=-\mathrm{E}\left[\psi'(Y)\cdot\left(\frac{\partial Y}{\partial S_0}\right)^2\right].$$

The LR method can also be applied because the transition density from S_i to S_{i+1} can be derived. We omit the derivation of the LR estimators here, because it is standard.

The WD method can also be applied. Interested readers are referred to the electronic companion for detailed derivation. However, it may require many additional simulations (e.g., up to $3k$ additional simulations[5] in this example). When the number of discretization steps k is relatively large, it may not be practical to apply the WD method.

To compare the performance of our approach to those of other methods, we use the estimated relative root mean-squared error (RRMSE) as a benchmark, which measures the percentage of the root mean-squared error to the absolute value of the quantity being estimated.

We vary k to examine the effect of the number of discretization steps on the estimation, and we vary the sample size n to examine the convergence of the estimators. RRMSEs of our pathwise approach (PW), SPA, the LR and WD methods are reported in Table 2. For delta and gamma, the WD method sometimes has smaller error than our approach. However, the WD method requires one and three additional simulations for delta and gamma, respectively, where the number in the bracket right after WD indicates how many additional simulations are required. Then in terms of efficiency, our approach outperforms the WD method. For vega and theta, the WD approach has larger errors and requires k and $3k$ additional simulations, respectively, and our approach is significantly better.

From the table we can also see that our approach is significantly better than SPA and the LR method, especially when k is large. The performance of our approach is stable with respect to the change of k, whereas the performances of SPA and the LR method become worse as k increases. By comparing the results for different n, we can also see that SPA and the LR method have faster convergence rates than our approach, which coincides with our analytical

Table 2. Estimating the Greeks of the Asian digital option.

	$n=10^3$			$n=10^4$			$n=10^5$		
k	10	50	100	10	50	100	10	50	100
RRMSE delta (%)									
PW	5.4	4.8	4.8	2.5	2.5	2.5	1.2	1.1	1.1
SPA	16.0	51.6	84.9	5.0	16.3	27.0	1.6	5.1	8.3
LR	13.2	24.7	32.3	4.2	7.8	10.5	1.3	2.3	3.3
WD (1)	6.4	12.0	17.6	1.9	3.8	5.3	0.6	1.2	1.8
RRMSE vega (%)									
PW	8.3	8.4	7.9	3.3	3.3	3.3	1.4	1.4	1.4
SPA	16.1	51.9	84.6	5.1	16.3	27.0	1.6	5.1	8.3
LR	23.0	48.3	67.0	7.1	15.6	21.2	2.2	4.8	6.6
WD (k)	44.7	221.2	437.6	14.2	67.0	135.7	4.4	22.4	46.6
RRMSE theta (%)									
PW	17.6	16.3	16.0	6.9	6.8	6.7	2.8	2.7	2.6
SPA	15.5	49.5	81.0	4.9	15.5	25.8	1.5	4.8	7.9
LR	22.0	45.1	62.4	7.0	14.4	19.6	2.1	4.5	6.2
WD (3k)	43.9	213.1	412.9	14.1	63.3	133.1	4.4	21.6	41.7
RRMSE gamma (%)									
PW	14.3	14.2	14.4	7.6	7.8	7.4	3.9	4.0	3.9
SPA	173	5,976	27,000	58.0	1,900	8,900	17.9	596	2,900
LR	36.4	134	245	11.6	39.9	80.4	3.6	13.0	24.5
WD (3)	16.7	77.0	158.9	5.5	25.2	48.9	1.7	8.1	15.2

Table 3. Approximately exact values of the Greeks of the barrier option.

	delta	vega	theta	gamma
$k=20$	-1.43×10^{-2}	-14.79	1.40	-7.49×10^{-3}
$k=50$	-1.88×10^{-2}	-14.37	1.37	-6.85×10^{-3}

result that convergence rate of our approach is $\sqrt{n\delta_n}$ for the first-order Greeks and $\sqrt{n\gamma_n^3}$ for the second-order Greeks, whereas that of SPA and the LR method is $\sqrt{n}$. However, even when n is as large as 10^5, our approach is still better than or comparable to SPA and the LR method. From the table we can also see that SPA and the LR method may not be appropriate for estimating gamma, whereas our approach works well.

4.2. Up-and-Out Barrier Call

Consider the up-and-out barrier call option introduced in §2.3.2. We suppose that S_t follows a geometric Brownian motion, i.e., $S_t = S_0 e^{(r-\sigma^2/2)t+\sigma B_t}$ where r and σ are risk-free interest rate and volatility, respectively, and B_t is a standard Brownian motion. In the numerical experiments we let $S_0 = K = 100$, $U = 120$, $r = 5\%$, $\sigma = 20\%$, and $T = 1$. There are no closed-form formulas for the Greeks in this example. However, by using finite difference with a large sample size (10^9), we find the approximately exact values of the Greeks, as summarized in Table 3 for different ks.

We compare the performance of our approach to those of the LR and WD methods, and the comparison results are summarized in Table 4. From the table we can see that our approach outperforms the LR method when the number of discretization steps k is large. Moreover, our approach has a smaller variance than the WD method for most Greeks

Table 4. Estimating the Greeks of the barrier call option.

	$n=10^4$		$n=10^5$		$n=10^6$	
k	20	50	20	50	20	50
RRMSE delta (%)						
PW	52.1	37.5	21.4	15.5	7.7	5.9
LR	64.4	74.7	19.5	22.8	6.2	7.3
WD (1)	33.7	38.5	10.9	12.3	3.4	4.0
RRMSE vega (%)						
PW	3.3	3.5	1.3	1.4	0.5	0.5
LR	8.6	13.7	2.8	4.5	0.9	1.4
WD (3k)	7.6	12.8	2.6	4.0	0.8	1.3
RRMSE theta (%)						
PW	4.8	5.9	2.2	2.4	0.8	1.0
LR	9.1	14.3	3.0	4.7	0.9	1.5
WD (3k)	8.1	13.5	2.8	4.2	0.9	1.4
RRMSE gamma (%)						
PW	21.6	24.4	10.2	11.3	4.8	5.7
LR	37.7	99.6	11.8	32.6	3.8	9.6
WD (4)	26.8	66.6	8.2	20.3	2.5	6.7

except delta. However, the WD method requires 1 resimulation for delta, and $3k$, $2k$, and 4 resimulations for vega, theta, and gamma, respectively. Then in terms of efficiency, our approach may outperform the WD method significantly.

5. Extension to the Sensitivity Analysis of Collateralized Debt Obligations

The proposed approach not only works for Greek estimation for equity options, but can also be applied to sensitivity analysis of many portfolio credit derivatives. In this example we consider an important class of portfolio credit derivatives, the so-called collateralized debt obligations (CDOs). We are interested in the sensitivities of the CDO value with respect to hazard rates of the underlying credit names, which provide important information for CDO hedging. For a comprehensive background on CDO and its sensitivities, readers are referred to Hull and White (2004) and Chen and Glasserman (2008) and the references therein.

The CDO contract being considered is associated with a portfolio of N underlying credit names that could be bonds, loans, or other defaultable assets. In this contract, credit loss of the portfolio is tranched to fractions, and different fractions of the loss are absorbed by investors with different risk preferences.

Following the notation in Chen and Glasserman (2008), we let T denote the life of the contract, τ_i the default time of the ith credit name, and l_i the loss given default of the ith credit name. Then the cumulative loss of the portfolio at time t (here $t \leqslant T$) is $L(t) = \sum_{i=1}^{N} l_i 1_{\{\tau_i \leqslant t\}}$, where τ_i follows an exponential distribution with mean $1/\lambda_i$, and the constant λ_i is called the hazard rate of the ith credit name.

For risk management purposes, one may be interested in the sensitivities of the CDO value with respect to the hazard rates of the underlying credit names. Without loss of generality, we consider its sensitivity with respect to λ_1. As shown in Chen and Glasserman (2008), the problem is then reduced to estimating

$$p'(\lambda_1) = \frac{\partial}{\partial \lambda_1} \mathrm{E}[(L(t) - y)^+]$$

for a given time $t \leqslant T$ and a given threshold value y.

Recall that $\tau_2, \ldots, \tau_N$ do not depend on λ_1. Then, applying Theorem 1, we have

$$p'(\lambda_1) = \frac{\partial}{\partial \lambda_1} \mathrm{E}\left[\left(\sum_{i=2}^{N} l_i 1_{\{\tau_i \leqslant t\}} + l_1 - y \right)^+ \cdot 1_{\{\tau_1 \leqslant t\}} \right.$$
$$\left. + \left(\sum_{i=2}^{N} l_i 1_{\{\tau_i \leqslant t\}} - y \right)^+ \cdot 1_{\{\tau_1 > t\}} \right]$$
$$= \partial_u \mathrm{E}\left[\left\{ \left(\sum_{i=2}^{N} l_i 1_{\{\tau_i \leqslant t\}} - y \right)^+ \right. \right.$$
$$\left. \left. - \left(\sum_{i=2}^{N} l_i 1_{\{\tau_i \leqslant t\}} + l_1 - y \right)^+ \right\} \frac{d\tau_1}{d\lambda_1} \cdot 1_{\{\tau_1 \leqslant t+u\}} \right]\Bigg|_{u=0}.$$

Table 5. Comparison of the estimators for CDO sensitivities.

β	$p'(\lambda_1)$	RRMSE (SPA) %	RRMSE (PW) %
0.3	0.290	1.71	3.25
0.2	0.284	2.23	3.26
0.1	0.278	3.20	3.21
0.05	0.278	4.70	3.32
0.01	0.278	10.81	3.33
0.005	0.278	15.42	3.27
0.001	0.278	34.70	3.28

Note that $d\tau_1/d\lambda_1 = \tau_1/\lambda_1$. Then we may apply the proposed kernel estimator to estimate $p'(\lambda_1)$.

As an illustrative example, we consider the case where the dependence among $\tau_1, \ldots, \tau_N$ is specified by a three-factor normal copula model. Interested readers are referred to §A.6 of the appendix for details and parameter specification of the model. Suppose that the CDO contract underlies $N = 200$ credit names that are divided into four groups, with hazard rates 0.5, 0.1, 0.12, 0.2, and losses given default 0.9, 0.6, 0.5, 0.1, respectively. We let $t = 0.75$ and $y = 15.75$, where y corresponds to 15% of the total losses of the portfolio.

Based on elaborate analysis, Chen and Glasserman (2008) propose an estimator of $p'(\lambda_1)$, which is essentially an SPA estimator. We compare our estimator (PW) to this SPA estimator for different scenarios. Specifically, we use an idiosyncratic risk control parameter β to indicate the level of the obligors' exposures to idiosyncratic risk, and we compare the estimators for different βs. When β is close to zero, the credit risk faced by an individual obligor is mainly caused by the systematic risk; when β is close to one, it is mainly caused by the idiosyncratic risk. We are particularly interested in cases where β is close to zero, because it is when many obligors default at the same time and the credit risk has the highest damage, as we have seen in the recent financial tsunami. For details of the idiosyncratic risk control parameter, readers are referred to A.6 of the appendix.

The comparison results are summarized in Table 5, where the sample size $n = 10^4$ and the bandwidth $\delta_n = c \times n^{-1/5}$ where $c = 0.8$ is obtained by the selection procedure provided in the electronic companion. From the table we can see that when β is large, the SPA estimator may have a smaller error and is better than our approach. However, when β becomes smaller and smaller, the error of our approach is rather stable while that of the SPA estimator blows up. Therefore, our approach may outperform the SPA estimator significantly when β is close to zero.

6. Conclusions

In this paper, we derive closed-form expressions for both the first- and second-order Greeks of options with discontinuous payoffs. Based on these expressions, we propose kernel estimators to estimate these Greeks. Our method

requires only the pathwise information that is observable from simulation and hence is easy to implement by practitioners. We also show that the approach proposed is valuable not only in risk management of equity options, but also in sensitivity analysis of portfolio credit derivatives.

7. Electronic Companion

An electronic companion to this paper is available as part of the online version that can be found at http://or.journal.informs.org/.

A. Appendix

A. 1. Proof of Lemma 1

PROOF. Note that

$$\partial_\theta \phi(\theta, y) = \partial_\theta E[g(S) \cdot 1_{\{h(S) \geqslant y\}}]$$
$$= \partial_\theta \int_{\Re^k} g(s) \cdot 1_{\{h(s) \geqslant y\}} f_{S(\theta)}(s)\, ds. \qquad (12)$$

Because for small-enough $\Delta\theta$,

$$|g(s) \cdot 1_{\{h(s) \geqslant y\}} f_{S(\theta+\Delta\theta)}(s) - g(s) \cdot 1_{\{h(s) \geqslant y\}} f_{S(\theta)}(s)|$$
$$\leqslant |g(s) \cdot 1_{\{h(s) \geqslant y\}}| \beta(s) \leqslant |g(s)| \beta(s),$$

and $\int_{\Re^k} |g(s)| \beta(s)\, ds \leqslant \infty$, by the dominated convergence theorem (Durrett 2005), we may take the partial differentiation inside the integral in the right-hand side of Equation (12), and hence

$$\partial_\theta \phi(\theta, y) = \int_{\Re^k} g(s) \cdot 1_{\{h(s) \geqslant y\}} \partial_\theta f_{S(\theta)}(s)\, ds.$$

Because $f_{S(\theta)}(s)$ is continuous differentiable in θ for almost all $s \in \Re^k$, and $g(S)$, $h(S)$, and $S(\theta)$ are continuous at θ almost surely for any $\theta \in \Theta$, it is easy to see that $g(s) \cdot 1_{\{h(s) \geqslant y\}} \partial_\theta f_{S(\theta)}(s)$ is continuous at $(\theta, 0)$ for almost all $s \in \Re^k$. Recalling further that $|f_{S(\theta+\Delta\theta)}(s) - f_{S(\theta)}(s)| \leqslant \beta(s)|\Delta\theta|$ for small-enough $\Delta\theta$, then $|\partial_\theta f_{S(\theta)}(s)| \leqslant \beta(s)$, and hence $\int_{\Re^k} g(s) \cdot 1_{\{h(s) \geqslant y\}} \partial_\theta f_{S(\theta)}(s)\, ds \leqslant \int_{\Re^k} |g(s)| \beta(s)\, ds < \infty$. Then by the dominated convergence theorem (Durrett 2005), we have

$$\lim_{(\theta', y) \to (\theta, 0)} \partial_\theta \phi(\theta', y) = \lim_{(\theta', y) \to (\theta, 0)} \int_{\Re^k} g(s) \cdot 1_{\{h(s) \geqslant y\}} \partial_\theta f_{S(\theta')}(s)\, ds$$

$$= \int_{\Re^k} \lim_{(\theta', y) \to (\theta, 0)} g(s) \cdot 1_{\{h(s) \geqslant y\}} \partial_\theta f_{S(\theta')}(s)\, ds$$

$$= \int_{\Re^k} g(s) \cdot 1_{\{h(s) \geqslant 0\}} \partial_\theta f_{S(\theta)}(s)\, ds$$

$$= \partial_\theta \phi(\theta, y),$$

which means $\partial_\theta \phi(\theta, y)$ is continuous at $(\theta, 0)$, and hence Assumption 2 is satisfied. $\square$

A.2. Proof of Theorem 1

PROOF. Let

$$\Pi(\theta, y) = g(S)[y - h(S)] \cdot 1_{\{h(S) \geqslant y\}} \quad \text{and} \quad \pi(\theta, y) = E[\Pi(\theta, y)].$$

Note that $\partial_y \Pi(\theta, y) = g(S) \cdot 1_{\{h(S) \geqslant y\}}$ w.p.1, and $|\Pi(\theta, y + \Delta y) - \Pi(\theta, y)| \leqslant |g(S)| \cdot |\Delta y|$ with $E[|g(S)|] < \infty$. Then, by the dominated convergence theorem (Durrett 2005),

$$\partial_y \pi(\theta, y) = E[\partial_y \Pi(\theta, y)] = E[g(S) \cdot 1_{\{h(S) \geqslant y\}}] = \phi(\theta, y).$$

Therefore,

$$
\begin{aligned}
p'(\theta) &= dE[g(S) \cdot 1_{\{h(S) \geqslant 0\}}]/d\theta = \partial_\theta \partial_y \pi(\theta, y)|_{y=0} \\
&= \partial_y \partial_\theta \pi(\theta, y)|_{y=0},
\end{aligned}
\tag{13}
$$

where the interchange of partial derivatives in the third equality follows from Assumption 2, that $\partial_\theta \partial_y \pi(\theta, y) = \partial_\theta \phi(\theta, y)$ is continuous at $(\theta, 0)$ (see, e.g., Marsden and Hoffman 1993).

Therefore, evaluating $p'(\theta)$ is equivalent to evaluating $\partial_y \partial_\theta \pi(\theta, y)|_{y=0}$. We first evaluate $\partial_\theta \pi(\theta, y)$. Note that

$$
\begin{aligned}
\partial_\theta \Pi(\theta, y) &= \partial_\theta g(S)[y - h(S)] \cdot 1_{\{h(S) \geqslant y\}} \\
&\quad - g(S) \partial_\theta h(S) \cdot 1_{\{h(S) \geqslant y\}} \quad \text{w.p.1.}
\end{aligned}
$$

Moreover, when $\Delta\theta$ is small enough,

$$
\begin{aligned}
&|\Pi(\theta + \Delta\theta, y) - \Pi(\theta, y)| \\
&\leqslant |g(S(\theta + \Delta\theta))[y - h(S(\theta + \Delta\theta))] \cdot 1_{\{g(S((\theta + \Delta\theta)) \geqslant y\}} \\
&\quad - g(S(\theta))[y - h(S(\theta + \Delta\theta))] \cdot 1_{\{h(S(\theta + \Delta\theta)) \geqslant y\}}| \\
&\quad + |g(S(\theta))[y - h(S(\theta + \Delta\theta))] \cdot 1_{\{h(S(\theta + \Delta\theta)) \geqslant y\}} \\
&\quad - g(S(\theta))[y - h(S(\theta))] \cdot 1_{\{h(S(\theta)) \geqslant y\}}| \\
&\leqslant |g(S(\theta + \Delta\theta)) - g(S(\theta))||y - h(S(\theta + \Delta\theta))| \\
&\quad + |g(S(\theta))| |[y - h(S(\theta + \Delta\theta))] \cdot 1_{\{h(S(\theta + \Delta\theta)) \geqslant y\}} \\
&\quad - [y - h(S(\theta))] \cdot 1_{\{h(S(\theta)) \geqslant y\}}| \\
&\leqslant (K_g |y - h(S(\theta + \Delta\theta))| + K_h |g(S(\theta))|) |\Delta\theta|.
\end{aligned}
$$

Note that

$$|h(S(\theta + \Delta\theta))| \leqslant |h(S(\theta))| + K_h |\Delta\theta| \leqslant |h(S(\theta))| + K_h$$

when $|\Delta\theta|$ is small enough. Then

$$
\begin{aligned}
&K_g |y - h(S(\theta + \Delta\theta))| + K_h |g(S(\theta))| \\
&\leqslant K_g^2 + y^2 + (|h(S(\theta))| + K_h)^2 + K_h^2 + |g(S(\theta))|^2,
\end{aligned}
$$

where the right-hand side of the equation has a finite mean by Assumption 1. Therefore, by the dominated convergence theorem (Durrett 2005),

$$
\begin{aligned}
\partial_\theta \pi(\theta, y) &= \partial_\theta E[\Pi(\theta, y)] = E[\partial_\theta \Pi(\theta, y)] \\
&= E[\partial_\theta g(S)(y - h(S)) \cdot 1_{\{h(S) \geqslant y\}}] \\
&\quad - E[g(S) \partial_\theta h(S) \cdot 1_{\{h(S) \geqslant y\}}].
\end{aligned}
\tag{14}
$$

We analyze the first term on the right-hand side of Equation (14). Because $\partial_\theta g(S)(y - h(S)) \cdot 1_{\{h(S) \geqslant y\}}$ is differentiable with respect to y w.p.1, and it is Lipschitz continuous in y with a Lipschitz constant $\partial_\theta g(S)$ which is bounded by K_g, then by the dominated convergence theorem (Durrett 2005),

$$
\begin{aligned}
&\partial_y E[\partial_\theta g(S)(y - h(S)) \cdot 1_{\{h(S) \geqslant y\}}] \\
&= E[\partial_y (\partial_\theta g(S)(y - h(S)) \cdot 1_{\{h(S) \geqslant y\}})] \\
&= E[\partial_\theta g(S) \cdot 1_{\{h(S) \geqslant y\}}].
\end{aligned}
\tag{15}
$$

Therefore, by Equations (13) to (15) we conclude the result of Theorem 1. $\square$

A.3. Proof of Theorem 2

PROOF. Let

$$\Pi(\theta_1, \theta_2, y) = g(S)[y - h(S)] \cdot 1_{\{h(S) \geqslant y\}} \quad \text{and}$$

$$\pi(\theta_1, \theta_2, y) = E[\Pi(\theta, y)].$$

Note that $\partial_y \pi(\theta_1, \theta_2, y) = \phi(\theta_1, \theta_2, y)$. Then, similar to the proof of Theorem 1, it is clear that

$$
\begin{aligned}
\partial_{\theta_2} \partial_{\theta_1} p(\theta_1, \theta_2) &= \partial_{\theta_2} \partial_{\theta_1} E[g(S) \cdot 1_{\{h(S) \geqslant 0\}}] \\
&= \partial_{\theta_2} \partial_{\theta_1} \partial_y \pi(\theta_1, \theta_2, y)|_{y=0} \\
&= \partial_y \partial_{\theta_2} \partial_{\theta_1} \pi(\theta, y)|_{y=0},
\end{aligned}
\tag{16}
$$

where the interchange of partial derivatives in the third equality follows from Assumption 4 that $\partial_{\theta_2} \partial_{\theta_1} \partial_y \pi(\theta_1, \theta_2, y) = \partial_{\theta_2} \partial_{\theta_1} \phi(\theta_1, \theta_2, y)$ is continuous at $(\theta, 0)$ (see, e.g., Marsden and Hoffman 1993).

By Equation (14), we have

$$
\begin{aligned}
\partial_{\theta_1} \pi(\theta_1, \theta_2, y) &= E[\partial_{\theta_1} g(S)(y - h(S)) \cdot 1_{\{h(S) \geqslant y\}}] \\
&\quad - E[g(S) \partial_{\theta_1} h(S) \cdot 1_{\{h(S) \geqslant y\}}].
\end{aligned}
\tag{17}
$$

With the same technique that is used to prove Equation (14), we can derive the partial derivative of the first term on the right-hand side of Equation (17),

$$
\begin{aligned}
&\partial_{\theta_2} E[\partial_{\theta_1} g(S)(y - h(S)) \cdot 1_{\{h(S) \geqslant y\}}] \\
&= E[\partial_{\theta_2} \partial_{\theta_1} g(S)(y - h(S)) \cdot 1_{\{h(S) \geqslant y\}}] \\
&\quad - E[\partial_{\theta_1} g(S) \partial_{\theta_2} h(S) \cdot 1_{\{h(S) \geqslant y\}}].
\end{aligned}
\tag{18}
$$

Applying Theorem 1 on the second term on the right-hand side of Equation (17), we have

$$
\begin{aligned}
&\partial_{\theta_2} E[g(S) \partial_{\theta_1} h(S) \cdot 1_{\{h(S) \geqslant y\}}] \\
&= E\{[\partial_{\theta_2} g(S) \partial_{\theta_1} h(S) + g(S) \partial_{\theta_2} \partial_{\theta_1} h(S)] \cdot 1_{\{h(S) \geqslant y\}}\} \\
&\quad - \partial_y E[g(S) \partial_{\theta_1} h(S) \partial_{\theta_2} h(S) \cdot 1_{\{h(S) \geqslant y\}}].
\end{aligned}
\tag{19}
$$

Therefore, combining Equations (17) to (19), we have

$$\partial_{\theta_2}\partial_{\theta_1}\pi(\theta_1,\theta_2,y) = E[\partial_{\theta_2}\partial_{\theta_1}h(S)[y-h(S)]\cdot 1_{\{h(S)\geqslant y\}}]$$
$$-E[(\partial_{\theta_1}g(S)\partial_{\theta_2}h(S)+\partial_{\theta_2}g(S)\partial_{\theta_1}h(S)$$
$$+g(S)\partial_{\theta_2}\partial_{\theta_1}h(S))\cdot 1_{\{\Lambda(\theta)\geqslant y\}}]$$
$$+\partial_y E[g(S)\partial_{\theta_1}h(S)\partial_{\theta_2}h(S)\cdot 1_{\{h(S)\geqslant y\}}].$$

Then, with the same techniques used in proving Equation (15), we have

$$\partial_y \partial_{\theta_2}\partial_{\theta_1}\pi(\theta_1,\theta_2,y)$$
$$=E[\partial_{\theta_2}\partial_{\theta_1}h(S)\cdot 1_{\{h(S)\geqslant y\}}]$$
$$-\partial_y E[(\partial_{\theta_1}g(S)\partial_{\theta_2}h(S)+\partial_{\theta_2}g(S)\partial_{\theta_1}h(S)$$
$$+g(S)\partial_{\theta_2}\partial_{\theta_1}h(S))\cdot 1_{\{\Lambda(\theta)\geqslant y\}}]$$
$$+\partial_y^2 E[g(S)\partial_{\theta_1}h(S)\partial_{\theta_2}h(S)\cdot 1_{\{h(S)\geqslant y\}}].$$

According to Equation (16), we obtain the conclusion of the theorem by setting $y=0$. □

A.4. Derivation of Equation (7)

Note that

$$\frac{dh(S)}{dS_0} = \frac{dS_k}{dS_0}1_{\{U-\check{S}>S_k-K\}} - \frac{d\check{S}}{dS_0}1_{\{S_k-K>U-\check{S}\}} \quad \text{w.p.1,} \quad (20)$$

where $d\check{S}/dS_0 = dS_{i^*}/dS_0$ and i^* is the index of the largest S_i among $S_1,\ldots,S_k$.

Then, by Theorem 1,

$$\frac{dp}{dS_0} = E\left[e^{-rT}\frac{dS_k}{dS_0}1_{\{S_k\geqslant K\}}1_{\{\check{S}\leqslant U\}}\right]$$
$$-\partial_y E\left[e^{-rT}(S_k-K)\frac{dS_k}{dS_0}1_{\{U-\check{S}>S_k-K\}}1_{\{S_k-K\geqslant y\}}\right]\Big|_{y=0}$$
$$+\partial_y E\left[e^{-rT}(S_k-K)\frac{d\check{S}}{dS_0}1_{\{S_k-K\geqslant U-\check{S}\}}1_{\{U-\check{S}\geqslant y\}}\right]\Big|_{y=0}.$$

By the definition of the differentiation, we can easily show that

$$\partial_y E\left[e^{-rT}(S_k-K)\frac{dS_k}{dS_0}1_{\{U-\check{S}>S_k-K\}}1_{\{S_k-K\geqslant y\}}\right]\Big|_{y=0} = 0,$$

and

$$\partial_y E\left[e^{-rT}(S_k-K)\frac{d\check{S}}{dS_0}1_{\{S_k-K\geqslant U-\check{S}\}}1_{\{U-\check{S}\geqslant y\}}\right]\Big|_{y=0}$$
$$=\partial_y E\left[e^{-rT}(S_k-K)\frac{d\check{S}}{dS_0}1_{\{S_k-K\geqslant 0\}}1_{\{U-\check{S}\geqslant y\}}\right]\Big|_{y=0}.$$

Then, assembling the terms, we obtain Equation (7).

A.5. Derivation of Equation (11)

We define

$$\xi_1 = E\left[e^{-rT}\frac{dS_k}{dS_0}1_{\{S_k\geqslant K\}}1_{\{\check{S}\leqslant U\}}\right],$$

$$\xi_2 = -\partial_y E\left[e^{-rT}(S_k-K)\frac{d\check{S}}{dS_0}1_{\{S_k-K\geqslant 0\}}1_{\{U-\check{S}\geqslant y\}}\right]\Big|_{y=0}.$$

Then, by Equation (7), $dp/dS_0 = \xi_1 - \xi_2$ and $d^2p/dS_0^2 = d\xi_1/dS_0 - d\xi_2/dS_0$.

First, we analyze $d\xi_1/dS_0$. Similar to the derivation of Equation (7),

$$\frac{d\xi_1}{dS_0} = E\left[e^{-rT}\frac{d^2S_k}{dS_0^2}1_{\{S_k\geqslant K\}}1_{\{\check{S}\leqslant U\}}\right]$$
$$-\partial_y E\left[e^{-rT}\left(\frac{dS_k}{dS_0}\right)^2 1_{\{U-\check{S}\geqslant 0\}}1_{\{S_k-K\geqslant y\}}\right]\Big|_{y=0}$$
$$+\partial_y E\left[e^{-rT}\frac{dS_k}{dS_0}\frac{d\check{S}}{dS_0}1_{\{S_k-K\geqslant 0\}}1_{\{U-\check{S}\geqslant y\}}\right]\Big|_{y=0}.$$

Second, we analyze $d\xi_2/dS_0$. Note that

$$\frac{d\xi_2}{dS_0} = -\partial_y \partial_{S_0}E\left[e^{-rT}(S_k-K)\frac{d\check{S}}{dS_0}1_{\{S_k-K\geqslant 0\}}1_{\{U-\check{S}\geqslant y\}}\right]\Big|_{y=0}.$$

By Theorem 1,

$$\partial_{S_0}E\left[e^{-rT}(S_k-K)\frac{d\check{S}}{dS_0}1_{\{S_k-K\geqslant 0\}}1_{\{U-\check{S}\geqslant y\}}\right]$$
$$=E\left[e^{-rT}\left(\frac{dS_k}{dS_0}\frac{d\check{S}}{dS_0}+(S_k-K)\frac{d^2\check{S}}{dS_0^2}\right)1_{\{S_k\geqslant K\}}1_{\{U-\check{S}\geqslant y\}}\right]$$
$$-\partial_u E\left[e^{-rT}(S_k-K)\left(\frac{d\check{S}}{dS_0}\right)^2 1_{\{S_k\geqslant K\}}1_{\{U-\check{S}\geqslant y+u\}}\right]\Big|_{u=0}.$$

Then,

$$\frac{d\xi_2}{dS_0} = -\partial_y E\left[e^{-rT}\left(\frac{dS_k}{dS_0}\frac{d\check{S}}{dS_0}+(S_k-K)\frac{d^2\check{S}}{dS_0^2}\right)\right.$$
$$\left.\cdot 1_{\{S_k\geqslant K\}}1_{\{U-\check{S}\geqslant y\}}\right]\Big|_{y=0}$$
$$+\partial_y^2 E\left[e^{-rT}(S_k-K)\left(\frac{d\check{S}}{dS_0}\right)^2 1_{\{S_k\geqslant K\}}1_{\{U-\check{S}\geqslant y\}}\right]\Big|_{y=0}.$$

Because $d^2p/dS_0^2 = d\xi_1/dS_0 - d\xi_2/dS_0$, then assembling the terms we obtain Equation (11).

A.6. Implementation Details of Normal Copula Model in §5

The dependence among $\tau_1,\ldots,\tau_N$ is specified in the following three-factor normal copula model. Specifically, $\tau_i =$

$F_i^{-1}(\Phi(W_i))$ where F_i is the exponential cumulative distribution function (C.D.F.) with mean $1/\lambda_i$, and Φ denotes the standard normal C.D.F. Here the latent variables $(W_1,\ldots,W_N)$ are generated by

$$W_k = a_{k1}Z_1 + a_{k2}Z_2 + a_{k3}Z_3 + b_k\epsilon_k, \quad \forall\, k = 1,\ldots,N,$$

where (Z_1, Z_2, Z_3) are systematic risk factors, each following a standard normal distribution, and ϵ_k is the kth obligor's idiosyncratic risk factor and follows a standard normal distribution. Here $A = (a_{ij})_{1 \leqslant i \leqslant N, 1 \leqslant j \leqslant 3}$ is the loading matrix with $a_{k1}^2 + a_{k2}^2 + a_{k3}^2 \leqslant 1$, and $b_k = \sqrt{1 - (a_{k1}^2 + a_{k2}^2 + a_{k3}^2)}$. The number b_k indicates the level of exposure of the kth obligor to its idiosyncratic risk.

To implement the model with $N = 200$ obligors, we need to specify the loading matrix A. To do this, we first randomly generate a base matrix A_0 whose entries $(\tilde{a}_{ij})$'s are independently and uniformly distributed over the interval $[-1/\sqrt{3}, 1/\sqrt{3}]$. Then for the first 10 rows of A_0, we multiply each row k by a factor $\sqrt{(1-\beta^2)/(\tilde{a}_{k1}^2 + \tilde{a}_{k2}^2 + \tilde{a}_{k3}^2)}.$, where β is an idiosyncratic risk control parameter. We then let this newly obtained matrix be the loading matrix A. By simple calculation we can see that with this loading matrix A the corresponding b_ks for the first 10 obligors are all β. Note that β being close to 0 implies that the first 10 obligors have small exposures to idiosyncratic risk, and hence their defaults are mainly caused by systematic risk.

Endnotes

1. The method is generally known as the pathwise method in the literature. In this paper, we generalize it and call the new one the pathwise method and the original one the classical pathwise method.
2. Here $\Delta\theta$ is independent of the realization of $S(\theta)$.
3. It was pointed out by an anonymous referee.
4. During implementation we use a pilot simulation with 500 sample paths. The computational time of the pilot simulation is very small and could be neglected. According to the numerical experiments, a relatively small sample size in the pilot simulation may lead to a reasonably good choice of δ_n.
5. The computation time of the WD method is roughly proportional to the number of total simulations required. For example, when the WD method requires $3k$ additional simulations, its computational time is roughly $3k+1$ times that of the pathwise method.

Acknowledgments

References

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

Bernis, G., E. Gobet, A. Kohatsu-Higa. 2003. Monte Carlo evaluation of Greeks for multidimensional barrier and lookback options. *Math. Finance* **13**(1) 99–113.

Bosq, D. 1998. *Nonparametric Statistics for Stochastic Processes*, 2nd ed. Springer, New York.

Broadie, M., P. Glasserman. 1996. Estimating security price derivatives using simulation. *Management Sci.* **42**(2) 269–285.

Chen, N., P. Glasserman. 2007. Malliavin Greeks without Malliavin calculus. *Stochastic Processes Their Appl.* **117**(11) 1689–1723.

Chen, N., L. J. Hong. 2007. Monte Carlo simulation in financial engineering. *Proc. 2007 Winter Simulation Conf.*, 919–931.

Chen, Z., P. Glasserman. 2008. Sensitivity estimates for portfolio credit derivatives using Monte Carlo. *Finance Stochastics* **12**(4) 507–540.

Durrett, R. 2005. *Probability: Theory and Examples*, 3rd ed. Duxbury Press, Belmont, CA.

Elie, R., J.-D. Fermanian, N. Touzi. 2007. Kernel estimation of Greek weights by parameter randomization. *Anal. Appl. Probab.* **14**(4) 1399–1423.

Fu, M. C. 2006. Gradient estimation. S. G. Henderson, B. L. Nelson, eds. *Simulation: Handbooks in Operations Research and Management Science.* Elsevier, Amsterdam, The Netherlands.

Fu, M., J.-Q. Hu. 1997. *Conditional Monte Carlo, Gradient Estimation and Optimization Applications.* Kluwer Academic Publishers, Boston.

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

Glynn, P. W. 1987. Likelihood ratio gradient estimation: An overview. *Proc. 1987 Winter Simulation Conf.*, 366–374.

Gong, W.-B., Y. C. Ho. 1987. Smoothed (conditional) perturbation analysis of discrete dynamical systems. *IEEE Trans. Automatic Control* **32**(10) 858–866.

Ho, Y. C., X. R. Cao. 1983. Perturbation analysis and optimization of queueing networks. *J. Optim. Theory Appl.* **40**(4) 559–582.

Hong, L. J. 2009. Estimating quantile sensitivities. *Oper. Res.* **57** 118–130.

Hong, L. J., G. Liu. 2010. Pathwise estimation of probability sensitivities through terminating or steady-state simulations. *Oper. Res.* **58**(2) 357–370.

Hull, J. C. 2006. *Options, Futures and Other Derivatives*, 6th ed. Pearson/Prentice Hall, Englewood Cliffs, NJ.

Hull, J. C., A. White. 2004. Valuation of a CDO and an nth default CDS without Monte Carlo simulation. *J. Derivatives* **12**(2) 8–23.

L'Ecuyer, P. 1991. An overview of derivative estimation. *Proc. 1991 Winter Simulation Conf.*, 207–217.

Marsden, J. E., M. J. Hoffman. 1993. *Elementary Classical Analysis*, 2nd ed. Freeman, New York.

Pflug, G. C. 1988. *Derivatives of Probability Measures—Concepts and Applications to the Optimization of Stochastic Systems.* Springer, Berlin.

Pflug, G. C., H. Weisshaupt. 2005. Probability gradient estimation by set-valued calculus and applications in network design. *SIAM J. Optim.* **15**(3) 898–914.