Efficient Price Sensitivity Estimation of Path-Dependent Derivatives by Weak Derivatives

Dr. Carlos Sanz Chacón[1]

[1]Goethe University Frankfurt
Computer Science and Mathematics
Numerical Analysis & Goethe Center for Scientific Computing (G-CSC)

6^{th} World Congress ot the Bachelier Finance Society
June 2010

1. Introduction
2. WDM: General Discussion
3. WDM: Models with Gaussian Transition Laws
4. Numerical Results
5. Summary

Motivation

- Development of more and more complicated financial products
  - more complex pricing
  - growing emphasis on risk management issues
  - global computation of risk figures such as VaR and CVaR
- Development of **efficient** methods for the computation of price sensitivities w.r.t. model parameters (“Greeks“)
  - Restriction: **computation time**, since, in many cases, these risk figures are not available in **closed formulas**
  - requirement of numerical methods!

Important Fact

Monte Carlo method is often the only applicable method!

Motivation

- Development of more and more complicated financial products
  - more complex pricing
  - growing emphasis on risk management issues
  - global computation of risk figures such as VaR and CVaR
- Development of **efficient** methods for the computation of price sensitivities w.r.t. model parameters ("Greeks")
  - Restriction: **computation time**, since, in many cases, these risk figures are not available in **closed formulas**
  - requirement of numerical methods!

Important Fact

Monte Carlo method is often the only applicable method!

Motivation

- Development of more and more complicated financial products
  - more complex pricing
  - growing emphasis on risk management issues
  - global computation of risk figures such as VaR and CVaR
- Development of **efficient** methods for the computation of price sensitivities w.r.t. model parameters (“Greeks“)
  - Restriction: **computation time**, since, in many cases, these risk figures are not available in **closed formulas**
  - requirement of numerical methods!

Important Fact

Monte Carlo method is often the only applicable method!

Description of the Estimation Problem

- **Derivative Price**: Let $J := \xi \cdot \mathbb{E}\big[L(Y)\big]$, where ξ denotes a deterministic discount factor, L a payoff and ϑ the parameter of interest. Then

$$\frac{d}{d\vartheta} J(\vartheta) = \xi(\vartheta) \cdot \frac{d}{d\vartheta} \mathbb{E}\big[L\big] + \mathbb{E}[L] \cdot \frac{d}{d\vartheta} \xi(\vartheta).$$

- **Complicated case**: Financial derivatives with discontinuous payoff L, e.g., $L := \mathbb{1}\{Y > K\}$
- **Mathematical problem**: Find a (random) vector-valued function g_ϑ such that

$$\nabla_\vartheta \mathbb{E}\big[L\big] = \mathbb{E}[g_\vartheta] \quad \Big(\nabla_\vartheta := \big(\tfrac{\partial}{\partial \vartheta_1}, \dots, \tfrac{\partial}{\partial \vartheta_n}\big)\Big),$$

where the function g_ϑ is called the stochastic gradient estimator of $\nabla_\vartheta \mathbb{E}\big[L\big]$.

Description of the Estimation Problem

- **Derivative Price**: Let $J := \xi \cdot \mathbb{E}\big[L(Y)\big]$, where ξ denotes a deterministic discount factor, L a payoff and ϑ the parameter of interest. Then

$$\frac{d}{d\vartheta} J(\vartheta) = \xi(\vartheta) \cdot \frac{\mathbf{d}}{\mathbf{d}\vartheta} \mathbb{E}\big[\mathbf{L}\big] + \mathbb{E}[L] \cdot \frac{d}{d\vartheta} \xi(\vartheta).$$

- **Complicated case**: Financial derivatives with discontinuous payoff L, e.g., $L := \mathbb{1}\{Y > K\}$
- **Mathematical problem**: Find a (random) vector-valued function g_ϑ such that

$$\nabla_\vartheta \mathbb{E}\big[L\big] = \mathbb{E}[g_\vartheta] \quad \Big(\nabla_\vartheta := \big(\tfrac{\partial}{\partial \vartheta_1}, \dots, \tfrac{\partial}{\partial \vartheta_n}\big)\Big),$$

where the function g_ϑ is called the stochastic gradient estimator of $\nabla_\vartheta \mathbb{E}\big[L\big]$.

Description of the Estimation Problem

- **Derivative Price**: Let $J := \xi \cdot \mathbb{E}\big[L(Y)\big]$, where ξ denotes a deterministic discount factor, L a payoff and ϑ the parameter of interest. Then

$$\frac{d}{d\vartheta} J(\vartheta) = \xi(\vartheta) \cdot \frac{\mathbf{d}}{\mathbf{d}\vartheta} \mathbb{E}\big[\mathbf{L}\big] + \mathbb{E}[L] \cdot \frac{d}{d\vartheta} \xi(\vartheta).$$

- **Complicated case**: Financial derivatives with discontinuous payoff L, e.g., $L := \mathbb{1}\{Y > K\}$
- **Mathematical problem**: Find a (random) vector-valued function g_ϑ such that

$$\nabla_\vartheta \mathbb{E}\big[L\big] = \mathbb{E}[g_\vartheta] \quad \Big(\nabla_\vartheta := \big(\tfrac{\partial}{\partial \vartheta_1}, \ldots, \tfrac{\partial}{\partial \vartheta_n}\big)\Big),$$

where the function g_ϑ is called the stochastic gradient estimator of $\nabla_\vartheta \mathbb{E}\big[L\big]$.

Standard Market Methods

- **Finite Difference (FD)**
  - $\frac{L(Y(\vartheta+\Delta\vartheta))-L(Y(\vartheta))}{\Delta\vartheta}$ (Forward FD Estimator)
  - Market standard (because very simple)
  - But biased estimator & large variance for discontinuous payoff L!
- **Score Function (SF)**
  - Introduced by Broadie & Glasserman (1996) to overcome disappointing performance of FD estimators
  - Estimator:

$$L(Y)\,\frac{d}{d\,\vartheta}\log f(Y;\vartheta),$$

    where f denotes the density of Y.
  - Benchmark for discontinuous payoff L, but the variance is still too large!

Standard Market Methods

- **Finite Difference (FD)**
  - $\frac{L(Y(\vartheta+\Delta\vartheta))-L(Y(\vartheta))}{\Delta\vartheta}$ (Forward FD Estimator)
  - Market standard (because very simple)
  - But biased estimator & large variance for discontinuous payoff L!
- **Score Function (SF)**
  - Introduced by Broadie & Glasserman (1996) to overcome disappointing performance of FD estimators
  - Estimator:

$$L(Y)\,\frac{d}{d\,\vartheta}\log f(Y;\vartheta),$$

    where f denotes the density of Y.
  - Benchmark for discontinuous payoff L, but the variance is still too large!

Weak Derivative Representation

- The WDM assumes the following representation:

$$\mathbb{E}_\vartheta[L(X)] = \int L(x)\,\mu_\vartheta(dx) \Rightarrow \frac{d}{d\vartheta}\,\mathbb{E}_\vartheta[L(X)] = \int L(x)\mu'_\vartheta(dx).$$

- **Main Idea**: Replace $\mu_\vartheta^{(k)}$ by one of the representations of its weak derivative.
- **One possibility: Hahn-Jordan decomposition**

$$\int L(x)\,\mu_\vartheta^{(k)}(dx) = c_\vartheta^{(k)}\left(\int L(x)\,\mu_\vartheta^{(k,+)}(dx) - \int L(x)\,\mu_\vartheta^{(k,-)}(dx)\right).$$

- Corresponding WD estimator: $\mathbf{g}_\vartheta^{(\mathbf{k})}(\mathbf{X}^{(\mathbf{k},+)}, \mathbf{X}^{(\mathbf{k},-)}) = \mathbf{c}_\vartheta^{(\mathbf{k})}(\mathbf{L}(\mathbf{X}^{(\mathbf{k},+)}) - \mathbf{L}(\mathbf{X}^{(\mathbf{k},-)}))$, where $X^{(k,+)} \sim \mu_\vartheta^{(k,+)}$ and $X^{(k,-)} \sim \mu_\vartheta^{(k,-)}$ are independent r.v.

Weak Derivative Representation

- The WDM assumes the following representation:

$$\mathbb{E}_\vartheta[L(X)] = \int L(x)\,\mu_\vartheta(dx) \Rightarrow \frac{d^{\mathbf{k}}}{d\vartheta^{\mathbf{k}}}\,\mathbb{E}_\vartheta[L(X)] = \int L(x)\mu_\vartheta^{(\mathbf{k})}(dx).$$

- **Main Idea**: Replace $\mu_\vartheta^{(k)}$ by one of the representations of its weak derivative.
- **One possibility: Hahn-Jordan decomposition**

$$\int L(x)\,\mu_\vartheta^{(k)}(dx) = c_\vartheta^{(k)}\left(\int L(x)\,\mu_\vartheta^{(k,+)}(dx) - \int L(x)\,\mu_\vartheta^{(k,-)}(dx)\right).$$

- Corresponding WD estimator:
$\mathbf{g}_\vartheta^{(\mathbf{k})}(\mathbf{X}^{(\mathbf{k},+)}, \mathbf{X}^{(\mathbf{k},-)}) = \mathbf{c}_\vartheta^{(\mathbf{k})}(\mathbf{L}(\mathbf{X}^{(\mathbf{k},+)}) - \mathbf{L}(\mathbf{X}^{(\mathbf{k},-)}))$, where $X^{(k,+)} \sim \mu_\vartheta^{(k,+)}$ and $X^{(k,-)} \sim \mu_\vartheta^{(k,-)}$ are independent r.v.

Weak Derivative Representation (cont'd)

- A stronger condition for weak differentiability is **absolute continuity**.
- This condition guarantees the existence of a density f_ϑ.
- The weak derivative of k^{th}-order has the representation

$$\frac{\partial^k f_\vartheta}{\partial \vartheta^k} = c_\vartheta^{(k)} \left(f_\vartheta^{(k,1)} - f_\vartheta^{(k,2)} \right),$$

where $f_\vartheta^{(k,1)}$ and $f_\vartheta^{(k,2)}$ are probability densities.

Model Setup

- System composed of a collection of independent normal r.v. $\{X_i; i = 1, \ldots, n\}$ with joint p.d.f.

$$\phi_\vartheta(x) := \prod_{i=1}^n \phi_{i,\vartheta}(x_i), \quad \text{where} \quad \phi_{i,\vartheta}(x_i) := \frac{e^{-\frac{1}{2}\left(\frac{x_i - \mu_i(\vartheta)}{\nu_i(\vartheta)}\right)^2}}{\sqrt{2\pi}\,\nu_i(\vartheta)}.$$

- Notice that this collection might describe a discrete Markov process $\{\widetilde{X}_i\}_{i=1}^n$ with deterministic initial value $\widetilde{X}_0 = \tilde{x}_0$ and transition p.d.f.

$$\phi_{i,\vartheta}(\tilde{x}_i; \tilde{x}_{i-1}) := \frac{e^{-\frac{1}{2}\left(\frac{\tilde{x}_i - \tilde{x}_{i-1} - \tilde{\mu}_i(\vartheta)}{\tilde{\nu}_i(\vartheta)}\right)^2}}{\sqrt{2\pi}\,\tilde{\nu}_i(\vartheta)} \quad \text{given} \quad \widetilde{X}_{i-1} = \tilde{x}_{i-1}.$$

Model Setup (cont'd)

- Notation:

$$X_{Y_i} := (X_1, \dots, X_{i-1}, Y_i, X_{i+1}, \dots, X_n),$$
$$X_{Y_{ij}} := (X_1, \dots, X_{i-1}, Y_i, X_{i+1}, \dots, X_{j-1}, Y_j, X_{j+1}, \dots, X_n) \quad (i \neq j),$$

  where Y_i and Y_j are independent r.v. with p.d.f. $f_{i,\vartheta}$ and $f_{j,\vartheta}$, respectively.
- Both r.v. are independent of all X_l $(l \neq i, j)$.
- The joint p.d.f. of the set of independent r.v. $\{X_1, \dots, Y_i, \dots, X_n\}$ and $\{X_1, \dots, Y_i, \dots, Y_j, \dots, X_n\}$ is given by

$$\prod_{l=1}^{i-1} \phi_{l,\vartheta}(x_l)\, f_{i,\vartheta}(y_i) \prod_{l=i+1}^{n} \phi_{l,\vartheta}(x_l) \quad \text{and} \quad \prod_{\substack{l=1 \\ l \neq i,j}}^{n} \phi_{l,\vartheta}(x_l)\, f_{i,\vartheta}(y_i)\; f_{j,\vartheta}(y_j),$$

  respectively.

Used Random Variables

- $X = (X_1, \ldots, X_n), \quad X_i \sim \mathrm{N}(\mu_i, \nu_i)$
- $W^{\pm} = (W_1^{\pm}, \ldots, W_n^{\pm}), \quad W_i^{\pm} \sim \mathrm{WB}(2, \pm\nu_i \sqrt{2}, \mu_i)$
  (WB = Weibull Distribution)
- $M = (M_1, \ldots, M_n), \quad M_i \sim \mathrm{DM}(\mu_i, \nu_i)$
  (DM = Double-Maxwell Distribution)
- $G^{\pm} = (G_1^{\pm}, \ldots, G_n^{\pm}), \quad G_i^{\pm} \sim \mathrm{G}(\mu_i, \pm\nu_i)$, with density
  $$f(x; \mu, \nu) := \begin{cases} \frac{1}{2\nu} \left(\frac{x-\mu}{\nu}\right)^3 \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\nu}\right)^2\right), & x \geq \mu \\ 0, & x < \mu \end{cases}$$
- $B = (B_1, \ldots, B_n), \quad B_i \sim \mathrm{B}(\mu_i, \nu_i)$ with density
  $$f(x; \mu, \nu) := \frac{1}{3\nu\sqrt{2\pi}} \left(\frac{x-\mu}{\nu}\right)^4 \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\nu}\right)^2\right), \quad \forall x \in \mathbb{R}$$

General WD estimator

Applying the WDM to $\frac{d}{d\vartheta}\xi(\vartheta)\,\mathbb{E}_\vartheta[L(X)]$ leads to the following unbiased sensitivity estimator $g_\vartheta^{(1)}$:

WD estimator for the 1^{st} derivative

$$g_\vartheta^{(1)}(x^{(1)}) = L(X)\frac{d}{d\vartheta}\xi(\vartheta) + \xi(\vartheta)\sum_{i=1}^{n}\left(a_{i,\vartheta}^{(1)}\,\Delta L_i^0 + b_{i,\vartheta}^{(1)}\,\Delta L_i^1\right)$$

$$x^{(1)} := (W^{\pm}, M, X),$$

$$\Delta L_i^0 := L(X_{W_i^+}) - L(X_{W_i^-}) \qquad \Delta L_i^1 := L(X_{M_i}) - L(X)$$

$$a_{i,\vartheta}^{(1)} := \frac{1}{\nu_i\sqrt{2\pi}}\,\frac{d\mu_i}{d\vartheta}, \qquad b_{i,\vartheta}^{(1)} := \frac{1}{\nu_i}\,\frac{d\nu_i}{d\vartheta}$$

General WD estimator (cont'd)

Applying the WDM to $\frac{d^2}{d\vartheta^2}\,\xi(\vartheta)\,\mathbb{E}_\vartheta[L(X)]$ leads to the following unbiased sensitivity estimator $g_\vartheta^{(2)}$:

WD estimator for the 2^{nd} derivative

$$g_\vartheta^{(2)}(x^{(2)}) = L(X)\frac{d^2}{d\vartheta^2}\xi(\vartheta) + \xi(\vartheta)\sum_{i=1}^{n}\Big\{a_{i,\vartheta}^{(2)}\,\Delta L_i^0 + b_{i,\vartheta}^{(2)}\,\Delta L_i^1 + c_{i,\vartheta}^{(2)}\big(2\,\Delta L_i^2 - 3\,\Delta L_i^0\big) + d_{i,\vartheta}^{(2)}\,\Delta L_i^1 + e_{i,\vartheta}^{(2)}\,\Delta L_i^3 + \sum_{j=1,\,j\neq i}^{n} a_{i,\vartheta}^{(k-1)}\;a_{j,\vartheta}^{(k-1)}\,\Delta L_{ij}^0 + b_{i,\vartheta}^{(k-1)}\;b_{j,\vartheta}^{(k-1)}\,\Delta L_{ij}^1\Big\} + 2\,\frac{d^{k-1}}{d\vartheta^{k-1}}\xi(\vartheta)\sum_{i=1}^{n} a_{i,\vartheta}^{(k-1)}\Delta L_i^0 + b_{i,\vartheta}^{(k-1)}\Delta L_i^1$$

General WD estimator (cont'd)

Applying the WDM to $\frac{d^2}{d\vartheta^2}\,\xi(\vartheta)\,\mathbb{E}_\vartheta[L(X)]$ leads to the following unbiased sensitivity estimator $g_\vartheta^{(2)}$:

WD estimator for the 2^{nd} derivative

$$g_\vartheta^{(2)}(x^{(2)}) = L(X)\frac{d^2}{d\vartheta^2}\xi(\vartheta) + \xi(\vartheta)\sum_{i=1}^{n}\Big\{ a_{i,\vartheta}^{(2)}\,\Delta L_i^0 + b_{i,\vartheta}^{(2)}\,\Delta L_i^1 + c_{i,\vartheta}^{(2)}\big(2\,\Delta L_i^2 - 3\,\Delta L_i^0\big) + d_{i,\vartheta}^{(2)}\,\Delta L_i^1 + e_{i,\vartheta}^{(2)}\,\Delta L_i^3 \Big\}$$

General WD estimator (cont'd)

WD estimator for the 2^{nd} derivative (cont'd)

$$x^{(2)} := (W^{\pm}, M, X, G^{\pm}, B),$$

$$\Delta L_i^2 := L(X_{G_i^+}) - L(X_{G_i^-}), \quad \Delta L_i^3 := 2\,L(X) + 3\,L(X_{B_i}) - 5\,L(X_{M_i}),$$

$$\Delta L_{ij}^0 := L(X_{W_{ij}^+}) - L(X_{W_{ij}^-}), \quad \Delta L_{ij}^1 := L(X_{M_{ij}}) - L(X),$$

$$a_{i,\vartheta}^{(2)} := \frac{1}{\nu_i\sqrt{2\pi}}\,\frac{d^2\mu_i}{d\vartheta^2}, \quad b_{i,\vartheta}^{(2)} := \left(\frac{1}{\nu_i}\,\frac{d\mu_i}{d\vartheta}\right)^2$$

$$c_{i,\vartheta}^{(2)} := \sqrt{\frac{2}{\pi}}\,\frac{1}{\nu_i^2}\,\frac{d\mu_i}{d\vartheta}\,\frac{d\nu_i}{d\vartheta}, \quad d_{i,\vartheta}^{(2)} := \frac{1}{\nu_i}\,\frac{d^2\nu_i}{d\vartheta^2},$$

$$e_{i,\vartheta}^{(2)} := \left(\frac{1}{\nu_i}\,\frac{d\nu_i}{d\vartheta}\right)^2.$$

General WD estimator (cont'd)

- $n_a := \max\{i \in \{1,\dots,n\} : a^{(1)}_{i,\vartheta} \neq 0\}$, $n_b := \max\{i \in \{1,\dots,n\} : b^{(1)}_{i,\vartheta} \neq 0\}$
- We make the following observations:

$$a^{(1)}_{i,\vartheta} = 0 \Rightarrow a^{(2)}_{i,\vartheta}, b^{(2)}_{i,\vartheta}, c^{(2)}_{i,\vartheta} = 0 \quad \text{and} \quad b^{(1)}_{i,\vartheta} = 0 \Rightarrow c^{(2)}_{i,\vartheta}, d^{(2)}_{i,\vartheta}, e^{(2)}_{i,\vartheta} = 0.$$

Reformulated WD estimators

$$g^{(1)}_{\vartheta}(x^{(1)}) = L(X)\frac{d}{d\vartheta}\xi(\vartheta) + \xi(\vartheta)\left(\sum_{i=1}^{n_a} a^{(1)}_{i,\vartheta}\,\Delta L^0_i + \sum_{i=1}^{n_b} b^{(1)}_{i,\vartheta}\,\Delta L^1_i\right)$$

$$g^{(2)}_{\vartheta}(x^{(2)}) = L(X)\frac{d^2}{d\vartheta^2}\xi(\vartheta) + \xi(\vartheta)\left[\sum_{i=1}^{n_a}\left(a^{(2)}_{i,\vartheta}\,\Delta L^0_i + b^{(2)}_{i,\vartheta}\,\Delta L^1_i\right)\right.$$
$$\left. + \sum_{i=1}^{n_b}\left(d^{(2)}_{i,\vartheta}\,\Delta L^1_i + e^{(2)}_{i,\vartheta}\,\Delta L^3_i\right) + \sum_{i=1}^{n_a \wedge n_b} c^{(2)}_{i,\vartheta}\left(2\,\Delta L^2_i - 3\,\Delta L^0_i\right)\right]$$

General WD estimator (cont‘d)

Remark

- If $n_a, n_b \ll n$, then the extra computational cost for the WD estimator is small compared to the SF estimator.
- The magnitude of n_a and n_b depend only on the concrete model and its model parameter ϑ.
- Fortunately, important price sensitivities of models used to price equity and FX derivatives have $n_a, n_b = 1$, e.g., Delta, Gamma and Theta. The BS and CEV models, considered here, are such models that have this advantageous property.

WD estimator in the BS model

WD estimator for Greeks

$$g_\vartheta^{(1)}(x^{(1)}) = (-1)^{q(\vartheta)}\, e^{-rT} \left[\sum_{i=1}^{n_a} \left(a_{i,\vartheta}^{(1)}\, \Delta L_i^0 + \sum_{i=1}^{n_b} b_{i,\vartheta}^{(1)}\, \Delta L_i^1 \right) - \frac{d\, rT}{d\vartheta}\, L(X) \right],$$

$$g_\vartheta^{(2)}(x^{(2)}) = (-1)^{q(\vartheta)}\, e^{-rT} \left[\sum_{i=1}^{n_a} \left(a_{i,\vartheta}^{(2)}\, \Delta L_i^0 + b_{i,\vartheta}^{(2)}\, \Delta L_i^1 \right) \right],$$

$$q(\vartheta) := \mathbb{1}\{\vartheta = T\},$$

$$a_{1,S_0}^{(1)} = \frac{1}{\nu\, S_0\, \sqrt{2\pi}}, \quad b_{1,S_0}^{(1)} = 0, \quad a_{i\geq 2,S_0}^{(1)} = 0, \quad b_{i\geq 2,S_0}^{(1)} = 0,$$

$$a_{1,S_0}^{(2)} = -\frac{1}{\nu\, S_0^2\, \sqrt{2\pi}}, \quad b_{1,S_0}^{(2)} = \frac{1}{\nu^2\, S_0^2}, \quad a_{i\geq 2,S_0}^{(2)} = 0, \quad b_{i\geq 2,S_0}^{(2)} = 0,$$

$$a_{i,\sigma}^{(1)} = -\sqrt{\frac{\Delta t}{2\,\pi}}, \quad b_{i,\sigma}^{(1)} = \frac{1}{\sigma}, \quad a_{i,r}^{(1)} = \frac{1}{\sigma}\sqrt{\frac{\Delta t}{2\,\pi}}, \quad b_{i,r}^{(1)} = 0,$$

$$a_{1,T}^{(1)} = \frac{r-\sigma^2/2}{\nu\,\sqrt{2\,\pi}}, \quad b_{1,T}^{(1)} = \frac{1}{2\,\Delta t}, \quad a_{i\geq 2,T}^{(1)} = 0, \quad b_{i\geq 2,T}^{(1)} = 0.$$

WD estimator in the CEV model

WD estimator for Greeks

$$g_\vartheta^{(1)}(x^{(1)}) = (-1)^{q(\vartheta)}\, e^{-rT} \left[\sum_{i=1}^{n_a}\Big(a_{i,\vartheta}^{(1)}\,\Delta L_i^0 + \sum_{i=1}^{n_b} b_{i,\vartheta}^{(1)}\,\Delta L_i^1\Big) - \tfrac{d\,rT}{d\vartheta}\,L(X)\right],$$

$$g_\vartheta^{(2)}(x^{(2)}) = (-1)^{q(\vartheta)}\, e^{-rT} \left[\sum_{i=1}^{n_a}\Big(a_{i,\vartheta}^{(2)}\,\Delta L_i^0 + b_{i,\vartheta}^{(2)}\,\Delta L_i^1\Big) + \sum_{i=1}^{n_b}\Big(d_{i,\vartheta}^{(2)}\,\Delta L_i^1 + e_{i,\vartheta}^{(2)}\,\Delta L_i^3\Big) + \sum_{i=1}^{n_a\wedge n_b} c_{i,\vartheta}^{(2)}\,\big(2\,\Delta L_i^2 - 3\,\Delta L_i^0\big)\right]$$

$$a_{1,S_0}^{(1)} = \frac{1}{\nu\, S_0\,\sqrt{2\pi}}, \qquad b_{1,S_0}^{(1)} = 0, \qquad a_{i\geq 2,S_0}^{(1)} = 0, \qquad b_{i\geq 2,S_0}^{(1)} = 0,$$

$$a_{1,S_0}^{(2)} = -\frac{1}{\nu\, S_0^2\,\sqrt{2\pi}}, \qquad b_{1,S_0}^{(2)} = \frac{1}{\nu^2\, S_0^2}, \qquad a_{i\geq 2,S_0}^{(2)} = 0, \qquad b_{i\geq 2,S_0}^{(2)} = 0,$$

$$a_{i,\sigma}^{(1)} = -\sqrt{\frac{\Delta t}{2\,\pi}}, \qquad b_{i,\sigma}^{(1)} = \frac{1}{\sigma}, \qquad a_{i,r}^{(1)} = \frac{1}{\sigma}\sqrt{\frac{\Delta t}{2\,\pi}}, \qquad b_{i,r}^{(1)} = 0,$$

$$a_{1,T}^{(1)} = \frac{r-\sigma^2/2}{\nu\,\sqrt{2\,\pi}}, \qquad b_{1,T}^{(1)} = \frac{1}{2\,\Delta t}, \qquad a_{i\geq 2,T}^{(1)} = 0, \qquad b_{i\geq 2,T}^{(1)} = 0.$$

Efficiency Measure

- Standard error (stderr): Precision of the mean estimate.
- Variance reduction factor VRF $= (\frac{\text{stderr}_{\text{BM}}}{\text{stderr}})^2$. Benchmark (BM) is SFM.
- Computational cost: Measured by the number of updates of the asset prices S_j's, $j = i, \ldots, n$ and $i = 1, \ldots, n^*$, $n^* \in \{n_a, n_b\}$.
  - The following ratio indicates a deterioration of performance:

$$\frac{\#\text{Calculation } S_j\text{'s}_{\text{original}} + \#\text{WDVariables} \times \#\text{Calculation } S_j\text{'s}_{\text{additional}}}{\#\text{Calculation } S_j\text{'s}_{\text{original}}}$$

  - #WDVariables denotes how many of the $Y \in \{W^{\pm}, M, G^{\pm}, B\}$ are involved in the WD estimator.
- Better efficiency measure for an estimator: *Divide the VRF by this ratio.*

AON Call: $L(S_T) = S_T \mathbb{1}\{S_T > K\}$

	FD	WD^U	WD
Δ:K=80	5.E-02	18	18
Δ:K=90	2.E-02	21	21
Δ:K=100	1.E-02	94	94
Δ:K=110	9.E-03	42	42
Δ:K=120	9.E-03	13	13
Δ:K=150	1.E-02	7	7
Γ:K=80	2.E-04	14	4
Γ:K=90	7.E-05	11	3
Γ:K=100	4.E-05	27	2
Γ:K=110	3.E-05	14	2
Γ:K=120	3.E-05	7	2
Γ:K=150	8.E-05	6	3

Table: VRF: BS Model

	FD	WD^U	WD
Δ:K=95	8.E-02	23	23
Δ:K=100	3.E-02	27	26
Δ:K=102	2.E-02	51	44
Δ:K=106	2.E-02	247	125
Δ:K=110	2.E-02	18	17
Δ:K=120	3.E-02	8	8
Γ:K=95	2.E-03	12	6
Γ:K=100	1.E-03	12	4
Γ:K=102	8.E-04	16	3
Γ:K=106	6.E-04	33	3
Γ:K=110	6.E-04	9	3
Γ:K=120	2.E-03	9	6

Table: VRF: CEV Model

AON Call: $L(S_T) = S_T \mathbb{1}\{S_T > K\}$

	FD	WD^U	WD
Δ:K=80	3.E-02	6	6
Δ:K=90	1.E-02	7	7
Δ:K=100	5.E-03	31	31
Δ:K=110	5.E-03	14	14
Δ:K=120	5.E-03	4	4
Δ:K=150	5.E-03	2	2
Γ:K=80	1.E-04	4	1
Γ:K=90	4.E-05	3	1
Γ:K=100	2.E-05	7	0.5
Γ:K=110	2.E-05	4	0.5
Γ:K=120	2.E-05	2	0.5
Γ:K=150	4.E-05	2	1

Table: VRF/Ratio: BS Model

	FD	WD^U	WD
Δ:K=95	4.E-02	6	6
Δ:K=100	2.E-02	7	7
Δ:K=102	1.E-02	13	11
Δ:K=106	1.E-02	62	31
Δ:K=110	1.E-02	5	4
Δ:K=120	2.E-02	2	2
Γ:K=95	1.E-03	2	1
Γ:K=100	5.E-04	2	1
Γ:K=102	4.E-04	2	0.5
Γ:K=106	3.E-04	5	0.5
Γ:K=110	3.E-04	1	0.5
Γ:K=120	1.E-03	1	1

Table: VRF/Ratio: CEV Model

Single Barrier AON call: $L(S) = S_T \mathbb{1}\{\min_{i=1,\dots,250} S_i > K\}$

	FD	WD^U	WD
Γ:K=80	3	125	55
Γ:K=85	2	65	29
Γ:K=90	0.8	32	15
Γ:K=95	0.3	13	6
Γ:K=100	0.1	7	2
Γ:K=102	0.2	5	3
κ:K=80	0.6	183	85
κ:K=85	0.4	112	51
κ:K=90	0.3	77	35
κ:K=95	0.2	56	25
κ:K=100	0.8	43	19
κ:K=102	0.1	36	17

Table: VRF: BS Model

	FD	WD^U	WD
Γ:K=90	102	261	118
Γ:K=95	19	49	22
Γ:K=97	9	23	11
Γ:K=98	6	14	7
Γ:K=100	2	7	2
Γ:K=101	4	6	4
κ:K=90	0.9	347	167
κ:K=95	0.3	106	50
κ:K=97	0.3	75	35
κ:K=98	0.2	61	30
κ:K=100	0.5	48	22
κ:K=101	0.05	35	19

Table: VRF: CEV Model

Single Barrier AON call: $L(S) = S_T \mathbb{1}\{\min_{i=1,\ldots,250} S_i > K\}$

	FD	WD^U	WD
Γ:K=80	2	31	14
Γ:K=85	1	16	7
Γ:K=90	0.4	8	4
Γ:K=95	0.2	3	2
Γ:K=100	0.05	2	0.5
Γ:K=102	0.05	1	1
κ:K=80	0.3	0.5	0.2
κ:K=85	0.2	0.3	0.1
κ:K=90	0.2	0.2	0.1
κ:K=95	0.1	0.1	0.1
κ:K=100	0.4	0.1	0.05
κ:K=102	0.05	0.1	0.05

Table: VRF/Ratio: BS Model

	FD	WD^U	WD
Γ:K=90	51	52	24
Γ:K=95	9	10	4
Γ:K=97	5	5	2
Γ:K=98	3	3	1
Γ:K=100	1	1	0.4
Γ:K=101	2	1	1
κ:K=90	0.5	3	1
κ:K=95	0.2	0.8	0.4
κ:K=97	0.2	0.6	0.3
κ:K=98	0.1	0.5	0.2
κ:K=100	0.3	0.4	0.2
κ:K=101	0.03	0.3	0.2

Table: VRF/Ratio: CEV Model

Fixed Lookback call: $L(S) = (\max_{i=1,\dots,250} S_i - K)\,\mathbb{1}\{\max_{i=1,\dots,250} S_i > K\}$

	FD	WD^U	WD
Γ:K=110	97	605	98
Γ:K=120	52	450	73
Γ:K=130	31	341	56
Γ:K=150	15	225	38
κ:K=110	473	824	137
κ:K=120	211	566	96
κ:K=130	102	425	71
κ:K=150	34	275	50

Table: VRF: BS Model

	FD	WD^U	WD
Γ:K=100	1778	1340	212
Γ:K=105	368	793	130
Γ:K=110	141	505	83
Γ:K=115	67	334	57
κ:K=100	1812	1918	307
κ:K=105	787	996	169
κ:K=110	255	592	101
κ:K=115	88	395	67

Table: VRF: CEV Model

Fixed Lookback call: $L(S) = (\max_{i=1,\dots,250} S_i - K)\,\mathbb{1}\{\max_{i=1,\dots,250} S_i > K\}$

	FD	WD^U	WD
Γ:K=110	49	151	25
Γ:K=120	26	113	18
Γ:K=130	16	85	14
Γ:K=150	8	56	10
κ:K=110	236	7	1
κ:K=120	106	4	0.8
κ:K=130	51	3	0.6
κ:K=150	17	2	0.4

Table: VRF/Ratio: BS Model

	FD	WD^U	WD
Γ:K=100	889	268	42
Γ:K=105	184	159	26
Γ:K=110	71	101	17
Γ:K=115	34	67	11
κ:K=100	906	15	2
κ:K=105	394	8	1
κ:K=110	128	5	0.8
κ:K=115	44	3	0.5

Table: VRF/Ratio: CEV Model

Summary

- Derivation of an unbiased WD sensitivity estimator in a Gaussian model framework
  - Valid for a large class of single-factor pricing models and path-dependent payoffs in use.
- From this general estimator we derived a WD estimator for all Greeks in the BS and CEV framework, respectively.
- Results of our simulation study
  - Coupled WD estimator had uniformly lower variance than the FD and SF estimator.
  - If the computational cost is taken into account, however, then only the Greeks with $n_a, n_b \ll n$, i.e., Δ, Γ and Θ, are more efficient than the standard methods. For κ and Θ this is not true any more.

Important Fact

WD estimator does not depend on the particular payoff but only on the underlying pricing model.

Summary

- Derivation of an unbiased WD sensitivity estimator in a Gaussian model framework
  - Valid for a large class of single-factor pricing models and path-dependent payoffs in use.
- From this general estimator we derived a WD estimator for all Greeks in the BS and CEV framework, respectively.
- Results of our simulation study
  - Coupled WD estimator had uniformly lower variance than the FD and SF estimator.
  - If the computational cost is taken into account, however, then only the Greeks with $n_a, n_b \ll n$, i.e., Δ, Γ and Θ, are more efficient than the standard methods. For κ and Θ this is not true any more.

Important Fact

WD estimator does not depend on the particular payoff but only on the underlying pricing model.

Summary

- Derivation of an unbiased WD sensitivity estimator in a Gaussian model framework
  - Valid for a large class of single-factor pricing models and path-dependent payoffs in use.
- From this general estimator we derived a WD estimator for all Greeks in the BS and CEV framework, respectively.
- Results of our simulation study
  - Coupled WD estimator had uniformly lower variance than the FD and SF estimator.
  - If the computational cost is taken into account, however, then only the Greeks with $n_a, n_b \ll n$, i.e., Δ, Γ and Θ, are more efficient than the standard methods. For κ and Θ this is not true any more.

Important Fact

WD estimator does not depend on the particular payoff but only on the underlying pricing model.

Summary

- Derivation of an unbiased WD sensitivity estimator in a Gaussian model framework
  - Valid for a large class of single-factor pricing models and path-dependent payoffs in use.
- From this general estimator we derived a WD estimator for all Greeks in the BS and CEV framework, respectively.
- Results of our simulation study
  - Coupled WD estimator had uniformly lower variance than the FD and SF estimator.
  - If the computational cost is taken into account, however, then only the Greeks with $n_a, n_b \ll n$, i.e., Δ, Γ and Θ, are more efficient than the standard methods. For κ and Θ this is not true any more.

Important Fact

WD estimator does not depend on the particular payoff but only on the underlying pricing model.

Authors of Related Literature

- **Price Sensitivity:**
  - Broadie, Glasserman (Pioneers in Estimation of Price Sensitivities)
  - Heidergott (Pricing of American Plain-Vanilla Call by Stochastic Optimization)
- **WDM:**
  - Pflug (Introduction of WDM)
  - Heidergott, Vazquez (Measure-Valued Differentiation)
  - Billingsley (Convergence of Probability Measures)
- **Overview:**
  - Fu (Summary of WDM and other approaches)

Analysis of Computational Cost

- In order to analyse the computational effort of the evaluation of L consider a stochastically recursive sequence (SRS)

$$S_l = h(S_{l-1}, X_l) \qquad l = 1, \dots, n, \tag{1}$$

  representing the underlying risk factor such as the asset price, where $S_0 = s_0$ is a deterministic start value, h is a measurable state-transition mapping and X_l denotes a random input variable distributed according to the Normal distribution.
- We point out that the nominal path $S(X)$ is generated from X, i.e.,

$$(X_1, \dots, X_n) \mapsto (S_1, \dots, S_n).$$

- It it obvious that all the S_l's are calculated and hence n calculations are needed.

Analysis of Computational Cost (cont'd)

- Note that the nominal path $S(X)$ and the perturbed path $S(X_{Y_i})$, $i = 1, \ldots, n^*$, are equal up to state S_{i-1} and differ from state S_i onwards, i.e., $n - i + 1$ states will change.
- This fact can be written as follows:

$$\begin{pmatrix} Y_1 & X_2 & \ldots & X_{n^*} & \ldots & X_n \\ X_1 & Y_2 & \ldots & X_{n^*} & \ldots & X_n \\ \vdots & \vdots & \ddots & \vdots & & \vdots \\ X_1 & X_2 & \ldots & Y_{n^*} & \ldots & X_n \end{pmatrix} \mapsto \begin{pmatrix} S_1^{Y_1} & S_2^{Y_1} & \ldots & S_{n^*}^{Y_1} & \ldots & S_n^{Y_1} \\ S_1 & S_2^{Y_2} & \ldots & S_{n^*}^{Y_2} & \ldots & S_n^{Y_2} \\ \vdots & \vdots & \ddots & \vdots & & \vdots \\ S_1 & S_2 & \ldots & S_{n^*}^{Y_{n^*}} & \ldots & S_n^{Y_{n^*}} \end{pmatrix}$$

- *How many of the S_j's, $j = i, \ldots, n$, need to be explicitly updated for the evaluation of the payoff L?*
- If we assume that all the S_j's need to be updated, then the answer is

$$\frac{1}{2} n^* (2n - n^* + 1).$$

Reduction Factors

	Table BS		Table CEV	
	FD	WD^U	FD	WD^U
Δ	2	3	2	4
Γ	2	4	2	7

Table: Reduction factors of VRF's for European AON call

	Table BS		Table CEV	
	FD	WD^U	FD	WD^U
Γ	2	4	2	5
κ	2	378	2	126

Table: Reduction factors of VRF's for European single barrier AON call

Reduction Factors (cont'd)

	Table BS		Table CEV	
	FD	WD^U	FD	WD^U
Γ	2	4	2	5
κ	2	378	2	126

Table: Reduction factors of VRF's for European fixed Lookback call

Weakly Differentiability

Definition (Weakly Differentiability)

Let $\vartheta, \vartheta + \Delta\vartheta \in V$. If μ_ϑ is a family of elements of $\mathcal{P}(\mathbb{R}^n)$, we say that μ_ϑ is weakly differentiable if there exists a finite signed measure $\mu_\vartheta^{'} : \mathfrak{B}(\mathbb{R}^n) \mapsto \mathbb{R}$ such that for all $L \in C_b(\mathbb{R}^n)$

$$\lim_{\Delta\vartheta \mapsto 0} \frac{1}{\Delta\vartheta}\left(\int L(x)\mu_{\vartheta+\Delta\vartheta}(dx) - \int L(x)\mu_\vartheta(dx) \right) = \int L(x)\mu_\vartheta^{'}(dx).$$

- Note that the above identity is for all $L \in C_b(\mathbb{R}^n)$ equivalent to

$$\frac{d}{d\vartheta} \int L(x)\mu_\vartheta(dx) = \int L(x)\mu_\vartheta^{'}(dx).$$

Representation of Weak Derivative

Definition (Representation of Weak Derivative)

Let μ_ϑ, $\vartheta \in V$, be a family of elements of $\mathcal{P}(\mathbb{R}^n)$. We call a triple $(c_\vartheta^{(k)}, \mu_\vartheta^{(k,1)}, \mu_\vartheta^{(k,2)})$ consisting of a constant and two probability measures a *representation of the weak derivative* of k^{th}-order of μ_ϑ if μ_ϑ is k-times weakly differentiable at each ϑ and for all $L \in C_b(\mathbb{R}^n)$ it holds that

$$\int L(x)\,\mu_\vartheta^{(k)}(dx) = c_\vartheta^{(k)}\left(\int L(x)\,\mu_\vartheta^{(k,1)}(dx) - \int L(x)\,\mu_\vartheta^{(k,2)}(dx)\right).$$

Remark to Discontinuous Payoffs

Remark

Assume that the sequence $\{\mu^{(k)}_{\Delta\vartheta} := \frac{\mu^{(k)}_{\vartheta+\Delta\vartheta} - \mu^{(k)}_{\vartheta}}{\Delta\vartheta} : \Delta\vartheta \in V \backslash \{0\}\}$ converges weakly to $\mu^{(k)}_{\vartheta}$. If for $L \in \mathcal{L}^1(\mu^{(k)}_{\Delta\vartheta} : \Delta\vartheta \in V \backslash \{0\})$ we denote by D_L the set of points at which L is discontinuous, then for each bounded L, such that $\mu^{(k)}_{\vartheta}(D_L) = 0$, the identity for the representation of weak derivatives is still true. For a proof of this result when the sequence $\{\mu^{(k)}_{\Delta\vartheta}\}_{\Delta\vartheta} \subset \mathcal{P}(\mathbb{R}^n)$ see, e.g., Billingsley (1999). The extension to the case $\{\mu^{(k)}_{\Delta\vartheta}\}_{\Delta\vartheta} \subset \mathcal{M}(\mathbb{R}^n)$ is straightforward to prove.

BS Model

- The asset price S_T is described in the BS model by

$$S_T = S_t \, e^{(r-\sigma^2/2)(T-t)+\sigma\sqrt{T-t}\,Z} \quad (0 \le t < T),$$

where $Z \sim \mathrm{N}(0,1)$, S_t is the current asset price, r and σ are constant.

- Divide the horizon $[0, T]$ into n equal time intervals each of length Δt: $0 = t_0 < t_1 < \cdots < t_n = T$.
- S_{t_i} at times t_i with known initial asset price S_0 are generated by

$$S_{t_i} = S_{t_{i-1}} \, e^{\tilde{\mu}+\tilde{\nu}\,Z_i} \quad (i = 1, \ldots, n),$$

where $Z_1, \ldots, Z_n$ is a sequence of independent, standard normal variables, $\tilde{\mu} := (r - \sigma^2/2)\,\Delta t$ and $\tilde{\nu} := \sigma\sqrt{\Delta t}$.

CEV Model

- Dynamics of asset price movements:

$$d\,S_t = r\,S_t\,dt + \sigma\,S_t^{\gamma}\,d\,W_t, \tag{2}$$

where W is a standard Brownian motion.

- The elasticity parameter γ was originally negative but was extended to include positive values.
- Euler approximation of (2)

$$S_{t_i} = S_{t_{i-1}} + \tilde{\mu}_i + \tilde{\nu}_i\,Z_i \quad (i = 1, \ldots, n), \tag{3}$$

where $\tilde{\mu}_i := r\,S_{t_{i-1}}\,\Delta t$ and $\tilde{\nu}_i := \sigma\,S_{t_{i-1}}^{\gamma}\sqrt{\Delta t}$.

Sensitivities

- Risk-neutral pricing formula:

$$V(0, S_0) = e^{-rT}\,\mathbb{E}\big[L(S)\big].$$

- Assumption (A0) is satisfied with ξ given by e^{-rT} in both models.
- The first ($k = 1$) and second derivative ($k = 2$) of the price w.r.t. the parameter ϑ is as follows:

$$\frac{d^k V}{d\vartheta^k} = -\frac{d^k\, rT}{d\vartheta^k}\, V + e^{-rT}\,\frac{d^k}{d\vartheta^k}\mathbb{E}[L] + \mathbb{1}\{k=2\}\left\{\left(\frac{d^{k-1}\, rT}{d\vartheta^{k-1}}\right)^2 V - 2\, e^{-rT}\,\frac{d^{k-1}\, rT}{d\vartheta^{k-1}}\,\frac{d^{k-1}}{d\vartheta^{k-1}}\mathbb{E}[L]\right\} \tag{4}$$

- In (4) $\frac{d^2\, rT}{d\vartheta^2} \equiv 0$ and $\frac{d\, rT}{d\vartheta}$ is only non-zero for $\vartheta \in \{r, T\}$.

The Greeks

Definition (The Greeks)

DELTA: The delta (Δ) is defined as the rate of change of the option price with respect to the initial asset price, i.e., $\Delta := \partial V / \partial S_0$.
GAMMA: The gamma (Γ) is the rate of change in the delta with respect to the initial asset price, i.e., $\Gamma := \partial^2 V / \partial S_0^2$.
VEGA: The vega (κ) is the rate of change of the option price with respect to the volatility of the underlying asset, i.e., $\kappa := \partial V / \partial \sigma$.
RHO: The rho (ρ) is defined as the rate of change of the option price with respect to the interest rate, i.e., $\rho := \partial V / \partial r$.
THETA: The theta (Θ) is the negative of the rate of change of the option price with respect to the passage of time, i.e., $\Theta := -\partial V / \partial T$.