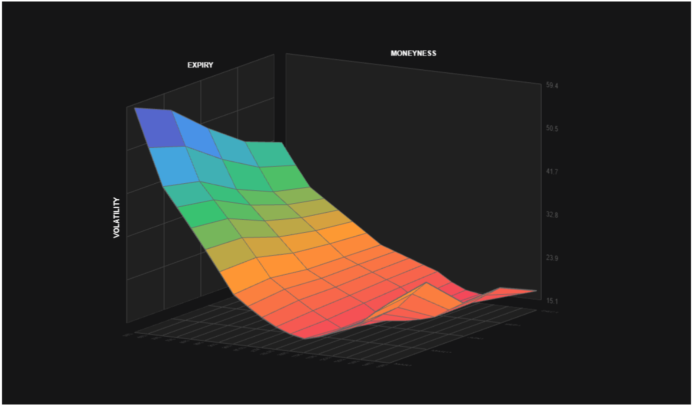
List of content

Deficiencies of the Black-Scholes Model

$\Rightarrow$ The idea of implied volatility does not fit to the Black-Scholes model

- Look for market consistent asset price models.

$\Rightarrow$ Use a local volatility, model stochastic volatility model, or a model with jumps, to better fit market data, and incorporate smile effects



Towards stochastic volatility

We have already seen the market:

$$\begin{cases} \mathrm{d}M(t) &= rM(t)\mathrm{d}t, \\ \mathrm{d}S(t) &= \mu S(t)\mathrm{d}t + \sigma S(t)\mathrm{d}W^{\mathbb{P}}(t), \end{cases}$$

where under $\mathbb{Q}$ measure $\mu = r$, i.e.:

$$\mathrm{d}S(t) = rS(t)\mathrm{d}t + \sigma S(t)\mathrm{d}W^{\mathbb{Q}}(t).$$

In the alternative process we aim to generalize the assumptions about constant parameters r and σ.

We can choose:

1. Constant: r, σ.
2. Deterministic- Piecewise constant: r_i, σ_i, on $[T_{i-1}, T_i]$.
3. Stochastic- time dependent: $r(t) = f(t, W_r(t))$, $\sigma(t) = g(t, W_\sigma(t))$.

Stochastic Volatility Models

- Modelling volatility as a random variable is confirmed by practical data that indicate the variable and unpredictable nature of volatility. (Hull and White, Stein and Stein, Heston, Schöbel and Zhu).
- The resulting SDE for the variance process can be recognized as a mean-reverting square-root process, a process originally proposed by Cox, Ingersoll & Ross (1985) to model the spot interest rate. If the variance exceeds its mean, it is driven back to the mean with the speed of mean reversion.
- Return distributions under stochastic volatility models also typically exhibit fatter tails than their log-normal counterparts, but the most significant argument to consider the volatility to be random is the implied volatility smile/skew, which can be accurately recovered by stochastic volatility models, especially for medium to long time to maturity options.

Heston Model

- The Heston model consists of two stochastic differential equations, for the underlying asset price, $S(t)$, and the variance process, $v(t)$, described under the risk-neutral measure, $\mathbb{Q}$, by

$$\begin{aligned} \mathrm{d}S(t) &= rS(t)\mathrm{d}t + \sqrt{v(t)}S(t)\mathrm{d}W_x(t), \\ \mathrm{d}v(t) &= \kappa(\bar{v} - v(t))\mathrm{d}t + \gamma\sqrt{v(t)}\mathrm{d}W_v(t). \end{aligned}$$

  Parameter interpretation.

- A correlation is defined between the underlying Brownian motions, $\mathrm{d}W_v(t)\mathrm{d}W_x(t) = \rho_{x,v}\mathrm{d}t$. Parameters $\kappa \geq 0$, $\bar{v} \geq 0$ and $\gamma > 0$ are called the speed of mean reversion, the long-term mean of the variance process and the volatility of the volatility, respectively.
  - r is the rate of the return,
  - $\bar{v}$ is the long vol, or long run average price volatility ($\lim_{t\to\infty} \mathbb{E}v(t) = \bar{v}$)
  - κ is the rate at which $v(t)$ reverts to $\bar{v}$,
  - γ is the vol- vol, or volatility of the volatility; as the name suggests, this determines the variance of $v(t)$.

Stochastic Volatility: Model of Heston

1. The variance process is a so-called CIR (Cox-Ingersoll-Ross) stochastic process:

$$\mathrm{d}v(t) \quad = \quad \kappa(\bar{v} - v(t))\mathrm{d}t + \gamma\sqrt{v(t)}\mathrm{d}W_v(t).$$

2. For a given time $t > 0$, variance $v(t)$ is distributed as $\bar{c}(t)$ times a noncentral chi-squared random variable, $\chi^2(\bar{d}, \bar{\lambda}(t))$, with $\bar{d}$ the "degrees of freedom" parameter and noncentrality parameter $\bar{\lambda}(t)$, i.e.

$$v(t) \sim \bar{c}(t)\chi^2\left(\bar{d}, \bar{\lambda}(t)\right), \quad t > 0,$$

with

$$\bar{c}(t) = \frac{1}{4\kappa}\gamma^2(1 - \mathrm{e}^{-\kappa t}), \quad \bar{d} = \frac{4\kappa\bar{v}}{\gamma^2}, \quad \bar{\lambda}(t) = \frac{4\kappa v_0 \mathrm{e}^{-\kappa t}}{\gamma^2(1 - \mathrm{e}^{-\kappa t})}.$$

3. The square-root process for the variance precludes negative values for $v(t)$, and if $v(t)$ reaches zero it can subsequently become positive. It is the Feller condition, $2\kappa\bar{v} \geq \gamma^2$, which guarantees that $v(t)$ stays positive; otherwise, if the Feller condition is not satisfied, the variance process may reach zero.

Noncentral χ^2-distribution

- Let $(X_1, X_2, \ldots, X_i, \ldots, X_{\bar{d}})$ be $\bar{d}$ independent, normally distributed random variables with means μ_i and variances σ_i^2. Then the random variable

$$\sum_{i=1}^{\bar{d}} \left(\frac{X_i}{\sigma_i}\right)^2$$

is distributed according to the noncentral chi-squared distribution.

- It has two parameters: $\bar{d}$ which specifies the number of degrees of freedom (i.e. the number of X_i), and noncentrality parameter $\bar{\lambda}(t)$ which is related to the mean of the random variables X_i by:

$$\bar{\lambda}(t) = \sum_{i=1}^{\bar{d}} \left(\frac{\mu_i}{\sigma_i}\right)^2.$$

- For this distribution we know the pdf, the characteristic function, the moment-generating function, etc.

Non-central Chi-squared distribution

- The corresponding cumulative distribution function (CDF):

$$F_{v(t)}(x) = P[v(t) \leq x] = P\left[\chi^2\left(\bar{d}, \bar{\lambda}(t)\right) \leq \frac{x}{\bar{c}(t)}\right] = F_{\chi^2(\bar{d},\bar{\lambda}(t))}\left(\frac{x}{\bar{c}(t)}\right),$$

where:

$$F_{\chi^2(\bar{d},\bar{\lambda}(t))}(y) = \sum_{k=0}^{\infty} \exp\left(-\frac{\bar{\lambda}(t)}{2}\right) \frac{\left(\frac{\bar{\lambda}(t)}{2}\right)^k}{k!} \frac{\Gamma\left(k + \frac{\bar{d}}{2}, \frac{y}{2}\right)}{\Gamma\left(k + \frac{\bar{d}}{2}\right)},$$

with $\Gamma(a, z) = \int_0^z t^{a-1}\mathrm{e}^{-t}\mathrm{d}t, \Gamma(z) = \int_0^\infty t^{z-1}\mathrm{e}^{-t}\mathrm{d}t.$

- The corresponding density function reads:

$$f_{\chi^2(\bar{d},\bar{\lambda}(t))}(y) = \frac{1}{2}\mathrm{e}^{-\frac{1}{2}(y+\bar{\lambda}(t))}\left(\frac{y}{\bar{\lambda}(t)}\right)^{\frac{1}{2}\left(\frac{\bar{d}}{2}-1\right)} \mathcal{B}_{\frac{\bar{d}}{2}-1}(\sqrt{\bar{\lambda}(t)y}),$$

with

$$\mathcal{B}_a(z) = \left(\frac{z}{2}\right)^a \sum_{k=0}^{\infty} \frac{\left(\frac{1}{4}z^2\right)^k}{k!\Gamma(a+k+1)},$$

PDF, CDF + Paths for CIR

- It is well-known that if the Feller condition, $2\kappa\bar{v} > \gamma^2$, is satisfied, the process $v(t)$ cannot reach zero, and if this condition does not hold the origin is accessible and strongly reflecting. In both cases, the $v(t)$ process cannot become negative.

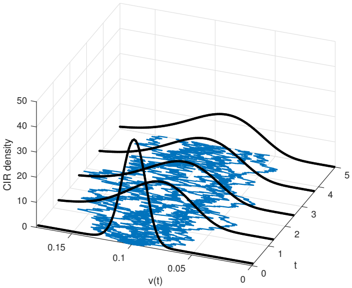
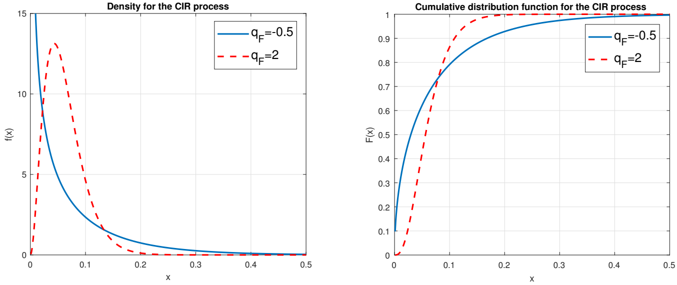


Figure: Paths and the corresponding PDF for the CIR process in the cases where the Feller condition is satisfied and is not satisfied. Simulations were performed with $\kappa = 0.5$, $v_0 = 0.1$, $\bar{v} = 0.1$. Left: $\gamma = 0.1$; Right: $\gamma = 0.35$.

PDF, CDF + Paths for CIR

- Feller condition is equivalent to "$\delta \geq 2$". By defining another parameter, $q_F := (2\kappa\bar{v}/\gamma^2) - 1$, the Feller condition is satisfied, when
$$q_F := \frac{2\kappa\bar{v}}{\gamma^2} - 1 = \frac{\delta}{2} - 1 \geq 0.$$
- There is one parameter set for which the Feller condition holds, i.e. $q_F = 2$, for $T = 5$, $\kappa = 0.5$, $v_0 = 0.2$, $\bar{v} = 0.05$, $\gamma = 0.129$ and one set for which the Feller condition is violated, $q_F = -0.5$, with $T = 5$, $\kappa = 0.5$, $v_0 = 0.2$, $\bar{v} = 0.05$, $\gamma = 0.316$.



Multi-dimensionality

- We need some mathematical tools for **multi-dimensional stochastic processes**.
- In the case of correlated Brownian motions, $\mathbb{E}[W_i(t) \cdot W_j(t)] = \rho_{i,j} t$, if $i \neq j$, and $\mathbb{E}[W_i(t) \cdot W_i(t)] = t$, if $i = j$, for $i, j = 1, \ldots n$.
- Similarly, for correlated Brownian increments, $\mathrm{d}W_i(t) \cdot \mathrm{d}W_j(t) = \rho_{i,j}\mathrm{d}t$, if $i \neq j$, and $\mathrm{d}W_i(t) \cdot \mathrm{d}W_i(t) = \mathrm{d}t$, if $i = j$.
- Two Brownian motions are independent, if $\mathbb{E}[\widetilde{W}_i(t) \cdot \widetilde{W}_j(t)] = 0$, if $i \neq j$ and $\mathbb{E}[\widetilde{W}_i(t) \cdot \widetilde{W}_j(t)] = t$, if $i = j$, for $i, j = 1, \ldots n$.
- For Brownian increments, $\mathrm{d}\widetilde{W}_i(t) \cdot \mathrm{d}\widetilde{W}_j(t) = 0$, if $i \neq j$ and $\mathrm{d}\widetilde{W}_i(t) \cdot \mathrm{d}\widetilde{W}_j(t) = \mathrm{d}t$, if $i = j$.

Cholesky decomposition; example

- Correlating two independent Brownian motions, $\widetilde{\mathbf{W}}(t) = [\widetilde{W}_1(t), \widetilde{W}_2(t)]^T$ with a correlation $\rho_{1,2}$.
- For a given (2×2)-correlation matrix $\mathbf{C}$, we find the Cholesky decomposition as

$$\mathbf{C} \stackrel{\text{def}}{=} \begin{bmatrix} 1 & \rho_{1,2} \\ \rho_{1,2} & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \rho_{1,2} & \sqrt{1-\rho_{1,2}^2} \end{bmatrix} \begin{bmatrix} 1 & \rho_{1,2} \\ 0 & \sqrt{1-\rho_{1,2}^2} \end{bmatrix}.$$

- To correlate independent Brownian motions, we calculate $\mathbf{L} \cdot \widetilde{\mathbf{W}}(t)$,

$$\begin{bmatrix} 1 & 0 \\ \rho_{1,2} & \sqrt{1-\rho_{1,2}^2} \end{bmatrix} \begin{bmatrix} \widetilde{W}_1(t) \\ \widetilde{W}_2(t) \end{bmatrix} = \begin{bmatrix} \widetilde{W}_1(t) \\ \rho_{1,2}\widetilde{W}_1(t) + \sqrt{1-\rho_{1,2}^2}\,\widetilde{W}_2(t) \end{bmatrix}.$$

Cholesky decomposition; example

- Defining $W_1(t) := \widetilde{W}_1(t)$, $W_2(t) := \rho_{1,2}\widetilde{W}_1(t) + \sqrt{1-\rho_{1,2}^2}\widetilde{W}_2(t)$, we determine the covariance between $W_1(t)$ and $W_2(t)$, as

$$
\begin{aligned}
\text{cov}[W_1(t), W_2(t)] &= \mathbb{E}[W_1(t)W_2(t)] - \mathbb{E}[W_1(t)]\mathbb{E}[W_2(t)] \\
&= \mathbb{E}\left[\widetilde{W}_1(t)\left(\rho_{1,2}\widetilde{W}_1(t) + \sqrt{1-\rho_{1,2}^2}\widetilde{W}_2(t)\right)\right] - 0 \\
&= \rho_{1,2}\mathbb{E}\left[(\widetilde{W}_1(t))^2\right] + \sqrt{1-\rho_{1,2}^2}\mathbb{E}[\widetilde{W}_1(t)\widetilde{W}_2(t)] \\
&= \rho_{1,2}\mathbb{E}\left[(\widetilde{W}_1(t))^2\right] + \sqrt{1-\rho_{1,2}^2}\mathbb{E}[\widetilde{W}_1(t)]\mathbb{E}[\widetilde{W}_2(t)] \\
&= \rho_{1,2}\mathbb{E}[(\widetilde{W}_1(t))^2] = \rho_{1,2}\mathrm{Var}[\widetilde{W}_1(t)] = \rho_{1,2}t.
\end{aligned}
$$

The correlation between $W_1(t)$ and $W_2(t)$ equals $\rho_{1,2}$, as desired.

Correlates Paths

In the first figure, the two Brownian motions are governed by a negative correlation parameter $\rho_{1,2}$, in the second figure $\rho_{1,2} = 0$, while in the third figure a positive correlation $\rho_{1,2} > 0$ is used.

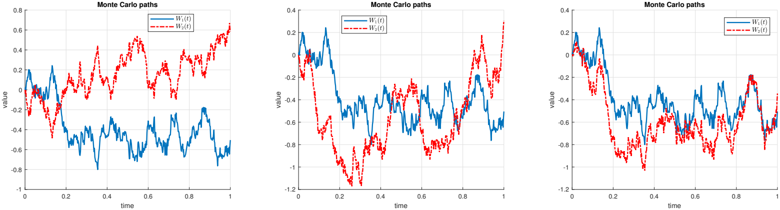


Figure: Monte Carlo Brownian motion $W_i(t)$ paths with different correlations, $\mathbb{E}[W_1(t)W_2(t)] = \rho_{1,2}t$; middle: zero correlation; left: negative correlation ($\rho_{1,2} < 0$), right: positive correlation ($\rho_{1,2} > 0$).

Cholesky decomposition

- The correlation structure is represented by matrix $\bar{\sigma}(t, \mathbf{X}(t))$. We then find:

$$\begin{aligned}\left(\mathbf{L}d\widetilde{\mathbf{W}}(t)\right)\left(\mathbf{L}d\widetilde{\mathbf{W}}(t)\right)^T &= (\mathbf{L}d\widetilde{\mathbf{W}}(t)d\widetilde{\mathbf{W}}(t)^T\mathbf{L}^T)\\ &= (\bar{\sigma}(t,\mathbf{X}(t))\bar{\sigma}(t,\mathbf{X}(t))^T)\cdot \text{diag}(\mathrm{d}t)\\ &=: \mathbf{C}\mathrm{d}t,\end{aligned}$$

  using $\mathrm{d}\widetilde{\mathbf{W}}(t)\mathrm{d}\widetilde{\mathbf{W}}(t)^T = \text{diag}(\mathrm{d}t)$.
- The correlation between the particular Brownian motions is represented by the instantaneous covariance matrix, $\bar{\sigma}(t, \mathbf{X}(t))$, via the Cholesky decomposition, as $\widetilde{\mathbf{W}}(t)$ are independent.
- Each symmetric positive definite matrix, $\mathbf{C}$, has a unique Cholesky decomposition, of the form, $\mathbf{C} = \mathbf{L}\mathbf{L}^T$, where $\mathbf{L}$ is a lower triangular matrix with positive diagonal entries.

Multi-dimensionality

- General system of *correlated SDEs*,

$$\mathrm{d}\mathbf{X}(t) = \bar{\boldsymbol{\mu}}(t, \mathbf{X}(t))\mathrm{d}t + \bar{\boldsymbol{\Sigma}}(t, \mathbf{X}(t))\mathrm{d}\mathbf{W}(t), \quad 0 \leq t_0 < t,$$

where $\bar{\boldsymbol{\mu}}(t, \mathbf{X}(t)) : D \to \mathbb{R}^n$, $\bar{\boldsymbol{\Sigma}}\,(t, \mathbf{X}(t)) : D \to \mathbb{R}^{n\times n}$ and $\mathbf{W}(t)$ is a column vector of correlated Brownian motions in $\mathbb{R}^n$.

- This SDE system can be written as:

$$\begin{aligned}
\begin{bmatrix} \mathrm{d}X_1 \\ \vdots \\ \mathrm{d}X_n \end{bmatrix} &= \begin{bmatrix} \bar{\mu}_1 \\ \vdots \\ \bar{\mu}_n \end{bmatrix}\mathrm{d}t + \begin{bmatrix} \bar{\Sigma}_{1,1} & \dots & \bar{\Sigma}_{1,n} \\ \vdots & \ddots & \vdots \\ \bar{\Sigma}_{n,1} & \dots & \bar{\Sigma}_{n,n} \end{bmatrix} \begin{bmatrix} \mathrm{d}W_1 \\ \vdots \\ \mathrm{d}W_n \end{bmatrix} \Leftrightarrow \\
\mathrm{d}\mathbf{X} &= \bar{\boldsymbol{\mu}}\mathrm{d}t + \bar{\boldsymbol{\Sigma}}\mathrm{d}\mathbf{W}.
\end{aligned}$$

Multi-dimensionality

- Using $\widetilde{\mathbf{W}}(t)$ is a column vector of *independent* Brownian motions in $\mathbb{R}^n$.
- With $\bar{\mu} = \bar{\mu}(t, \mathbf{X}(t))$, $\bar{\sigma} = \bar{\sigma}(t, \mathbf{X}(t))$ and $\widetilde{W} = \widetilde{W}(t)$, the dynamics for $\mathbf{X} = \mathbf{X}(t)$ give the matrix representation:

$$
\begin{aligned}
\begin{bmatrix} \mathrm{d}X_1 \\ \vdots \\ \mathrm{d}X_n \end{bmatrix} &= \begin{bmatrix} \bar{\mu}_1 \\ \vdots \\ \bar{\mu}_n \end{bmatrix} \mathrm{d}t + \begin{bmatrix} \bar{\sigma}_{1,1} & \dots & \bar{\sigma}_{1,n} \\ \vdots & \ddots & \vdots \\ \bar{\sigma}_{n,1} & \dots & \bar{\sigma}_{n,n} \end{bmatrix} \begin{bmatrix} \mathrm{d}\widetilde{W}_1 \\ \vdots \\ \mathrm{d}\widetilde{W}_n \end{bmatrix} \\
&= \bar{\mu}\mathrm{d}t + \bar{\mathbf{\Sigma}}\mathbf{L}\mathrm{d}\widetilde{\mathbf{W}} = \bar{\mu}\mathrm{d}t + \bar{\boldsymbol{\sigma}}\mathrm{d}\widetilde{\mathbf{W}}.
\end{aligned}
$$

Itô's lemma for vector processes

- Consider $\mathbf{X}(t) = [X_1(t), X_2(t), \ldots, X_n(t)]^T$ and let a real-valued function $g \equiv g(t, \mathbf{X}(t))$ be sufficiently differentiable on $\mathbb{R} \times \mathbb{R}^n$. Increment $dg(t, \mathbf{X}(t))$ is governed by the following SDE:

$$\mathrm{d}g(t, \mathbf{X}(t)) = \frac{\partial g}{\partial t}\mathrm{d}t + \sum_{j=1}^{n} \frac{\partial g}{\partial X_j}\mathrm{d}X_j(t) + \frac{1}{2}\sum_{i,j=1}^{n} \frac{\partial^2 g}{\partial X_i \partial X_j}\mathrm{d}X_i(t)\mathrm{d}X_j(t).$$

- Using the matrix notation, we distinguish the drift and the volatility terms in $\mathrm{d}g(t, \mathbf{X}(\mathbf{t})) =$

$$\left(\frac{\partial g}{\partial t} + \sum_{i=1}^{n} \bar{\mu}_i(t, \mathbf{X}(t))\frac{\partial g}{\partial X_i} + \frac{1}{2}\sum_{i,j,k=1}^{n} \bar{\sigma}_{i,k}(t, \mathbf{X}(t))\bar{\sigma}_{j,k}(t, \mathbf{X}(t))\frac{\partial^2 g}{\partial X_i \partial X_j}\right)\mathrm{d}t$$
$$+ \sum_{i,j=1}^{n} \bar{\sigma}_{i,j}(t, \mathbf{X}(t))\frac{\partial g}{\partial X_i}\mathrm{d}\widetilde{W}_j(t).$$

  This is found by application of Taylor series expansion, and the Itô table.

Example: 2D correlated geometric Brownian motion

- With 2D Brownian motion $\mathbf{W}(t) = [W_1(t), W_2(t)]^T$, and correlation parameter ρ, we construct a portfolio consisting of two correlated stocks, $S_1(t)$ and $S_2(t)$, with dynamics:

$$\begin{aligned} \mathrm{d}S_1(t) &= \mu_1 S_1(t)\mathrm{d}t + \sigma_1 S_1(t)\mathrm{d}W_1(t), \\ \mathrm{d}S_2(t) &= \mu_2 S_2(t)\mathrm{d}t + \sigma_2 S_2(t)\mathrm{d}W_2(t), \end{aligned}$$

  with $\mu_1, \mu_2, \sigma_1, \sigma_2$ constants.
- By the Cholesky decomposition this system can be expressed, in terms of independent Brownian motions, as:

$$\left[\begin{array}{c} \mathrm{d}S_1(t) \\ \mathrm{d}S_2(t) \end{array}\right] = \left[\begin{array}{c} \mu_1 S_1(t) \\ \mu_2 S_2(t) \end{array}\right]\mathrm{d}t + \left[\begin{array}{cc} \sigma_1 S_1(t) & 0 \\ \rho\sigma_2 S_2(t) & \sqrt{1-\rho^2}\sigma_2 S_2(t) \end{array}\right]\left[\begin{array}{c} \mathrm{d}\widetilde{W}_1(t) \\ \mathrm{d}\widetilde{W}_2(t) \end{array}\right].$$

Example: 2D correlated geometric Brownian motion

- Application of multi-D Itô lemma to a sufficiently smooth function, $g \equiv g(t, S_1, S_2)$, $S_i = S_i(t)$, $i = 1, 2$, gives:

$$\mathrm{d}g(t, S_1, S_2) = \left(\frac{\partial g}{\partial t} + \mu_1 S_1 \frac{\partial g}{\partial S_1} + \mu_2 S_2 \frac{\partial g}{\partial S_2} + \frac{1}{2}\sigma_1^2 S_1^2 \frac{\partial^2 g}{\partial S_1^2} + \frac{1}{2}\sigma_2^2 S_2^2 \frac{\partial^2 g}{\partial S_2^2} \right.$$
$$\left. + \rho\sigma_1\sigma_2 S_1 S_2 \frac{\partial^2 g}{\partial S_1 \partial S_2}\right) \mathrm{d}t + \sigma_1 S_1 \frac{\partial g}{\partial S_1} \mathrm{d}W_1 + \sigma_2 S_2 \frac{\partial g}{\partial S_2} \mathrm{d}W_2.$$

- This result holds for any function $g(t, S_1, S_2)$ which satisfies the differentiability conditions.
- If we, for example, take $g(t, S_1, S_2) \equiv \log S_1$ the result collapses to the well-known dynamics for the log-stock:

$$\mathrm{d}\log S_1(t) = \left(\mu_1 - \frac{1}{2}\sigma_1^2\right) \mathrm{d}t + \sigma_1 \mathrm{d}W_1(t).$$

Back to Heston stoch. vol. dynamics

- The *martingale method* can be used to determine option pricing PDE under Heston dynamics.
- For Heston's model, we consider the following pricing problem:

$$V(S,v,t) = M(t)\mathbb{E}^{\mathbb{Q}}\left[\frac{1}{M(T)}V(S,v,T)\Big|\mathcal{F}(t)\right], \tag{1}$$

Under the usual regularity assumptions, we assume the existence of a differentiable function, $\Pi_V \equiv \Pi_V(S,v,t)$, which is a martingale,

$$\mathbb{E}^{\mathbb{Q}}\left[\frac{1}{M(T)}V(S,v,T)\Big|\mathcal{F}(t)\right] = \frac{V(S,v,t)}{M(t)} =: \Pi_V(S,v,t). \tag{2}$$

- By the martingale definition, we can determine the dynamics using Itô's lemma,

$$\mathrm{d}\Pi_V = \mathrm{d}\left(\frac{V}{M}\right) = \frac{1}{M}\mathrm{d}V - r\frac{V}{M}\mathrm{d}t.$$

- An infinitesimal change, $\mathrm{d}V(S,v,t)$, with the dynamics for $S(t)$ and $v(t)$ in the Heston model, gives

$$\begin{aligned}\mathrm{d}V &= \left(\frac{\partial V}{\partial t} + rS\frac{\partial V}{\partial S} + \kappa(\bar{v}-v)\frac{\partial V}{\partial v} + \frac{1}{2}vS^2\frac{\partial^2 V}{\partial S^2}\right.\\ &+ \left.\rho_{x,v}\gamma S v\frac{\partial^2 V}{\partial S\partial v} + \frac{1}{2}\gamma^2 v\frac{\partial^2 V}{\partial v^2}\right)\mathrm{d}t\\ &+ S\sqrt{v}\frac{\partial V}{\partial S}\mathrm{d}W_x + \gamma\sqrt{v}\frac{\partial V}{\partial v}\mathrm{d}W_v.\end{aligned}$$

- $\mathrm{d}\Pi_V(S,v,t)$ should be free of $\mathrm{d}t$-terms:

$$\frac{1}{M}\left(\frac{\partial V}{\partial t} + rS\frac{\partial V}{\partial S} + \kappa(\bar{v}-v)\frac{\partial V}{\partial v} + \frac{1}{2}vS^2\frac{\partial^2 V}{\partial S^2} + \rho_{x,v}\gamma S v\frac{\partial^2 V}{\partial S\partial v} + \frac{1}{2}\gamma^2 v\frac{\partial^2 V}{\partial v^2}\right) - r\frac{V}{M} = 0,$$

resulting in the option pricing PDE for the Heston model.

Interpretation of Model Parameters

- In the Black-Scholes model the variance, σ^2, is constant in the Heston model it is driven by a mean-reverting stochastic process, $v(t)$,
$$\mathrm{d}v(t) = \kappa(\bar{v} - v(t))\mathrm{d}t + \gamma\sqrt{v(t)}\mathrm{d}W_v(t).$$
- In the Heston model each parameter has a specific effect on the implied volatility curve generated by the dynamics,
- To analyze the parameter effects numerically with an example, we use here the following set of reference parameters, $\rho_{x,v} = 0\%$, $\kappa = 1$, $\gamma = 0.1$, $v_0 = 0.05$ and $\bar{v} = 0.1$.
- We change individual parameters while keeping the others fixed. For each parameter set option prices have been generated and inserted in a Newton-Raphson algorithm to determine the implied volatilities.

- Correlation, $\rho_{x,v}$, and vol-vol parameter, γ. For $\rho_{x,v} = 0\%$ a higher value of γ gives a more pronounced implied volatility *smile*. A higher volatility-of-volatility parameter increases the implied volatility curvature.
- As the correlation between stock and variance process, $\rho_{x,v}$, gets increasingly negative the slope of the skew in the implied volatility curve increases.

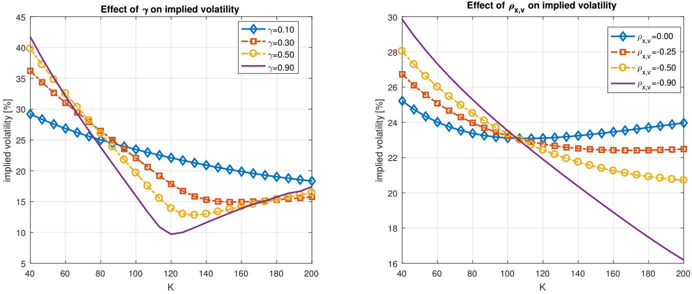


Figure: Impact of variation of the Heston vol-vol parameter γ (left side), and correlation parameter $\rho_{x,v}$ (right side), on the implied volatility, as a function of strike price K.

- Speed of mean reversion κ has a limited effect on the implied volatility smile or skew, up to $1\% - 2\%$. κ determines the speed at which the volatility converges to the long-term volatility $\bar{v}$, see the RHS graph, which shows the at-the-money (ATM) implied volatility for different κ.
- With $\bar{v} = 10\%$ ($\sqrt{\bar{v}} \approx 31.62\%$) a large κ-value implies fast convergence of the implied volatility to $\sqrt{\bar{v}}$.

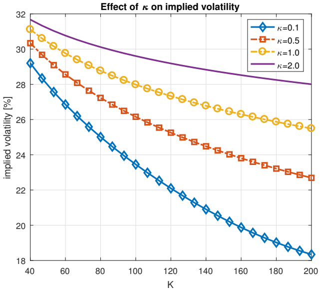


Figure: Impact of variation of the Heston parameter κ on the implied volatility as a function of strike K (left side), impact of variation of κ on the ATM volatility, as a function of $\tau = T - t$ (right side).

- v_0, the initial variance and $\bar{v}$, the variance level, have a similar effect on the implied volatility curve.
- The effect of these two parameters seems to depend on the value of κ, controlling the speed at which the implied volatility converges from $\sqrt{v_0}$ to $\sqrt{\bar{v}}$ (or v_0 to $\bar{v}$).

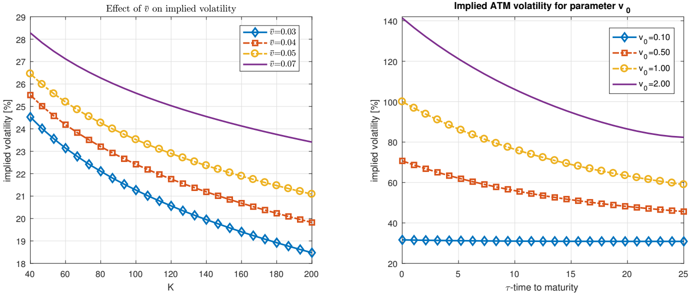


Figure: Impact of changing v_0 and $\bar{v}$ on the Heston implied volatility; left side: $\bar{v}$ as a function of the strike K, right side: v_0 as a function of time to maturity $\tau = T - t$.

Black-Scholes vs. Heston

Set $T = 2$; $v_0 = 0.1$; $r = 0.05$; $S_0 = 1$; $\kappa = 0.2$; $\bar{v} = 0.3$; $\gamma = 0.1$; $\rho = -0.8$;

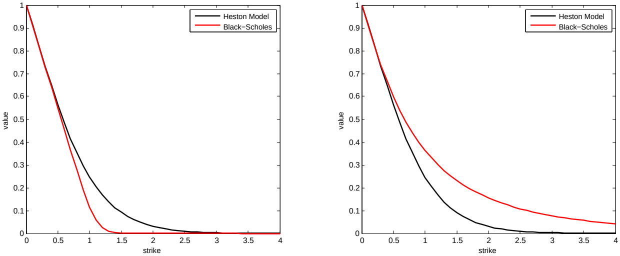


Figure: LEFT: $\sigma^{BS} = \sqrt{v_0}$, RIGHT: $\sigma^{BS} = 60\%$

- An inspection of Heston's model does reveal some important differences with respect to GBM.
- The probability density functions of (log-)returns have heavier tails, compared to Gaussian;
- The volatility smile can be represented by parameter combinations
- The following option pricing PDE under the Heston dynamics:

$$\frac{\partial V}{\partial t} + \frac{1}{2} v S^2 \frac{\partial^2 V}{\partial S^2} + \rho_{x,v} \gamma S v \frac{\partial^2 V}{\partial S \partial v} + \frac{1}{2} \gamma^2 v \frac{\partial^2 V}{\partial v^2} + rS \frac{\partial V}{\partial S} + \kappa(\bar{v} - v(t)) \frac{\partial V}{\partial v} - rV = 0.$$

Heston Model

From the definition of the Heston model we have:

$$\begin{cases} \mathrm{d}S(t) &= rS(t)\mathrm{d}t + \sqrt{v(t)}S(t)\mathrm{d}W_x(t) \\ \mathrm{d}v(t) &= \kappa\left(\bar{v} - v(t)\right)\mathrm{d}t + \gamma\sqrt{v(t)}\mathrm{d}W_v(t) \end{cases}$$

Is it affine?

$$\sigma(\mathbf{X}(\mathbf{t}))\sigma(\mathbf{X}(\mathbf{t}))^T = \begin{bmatrix} v(t)S(t)^2 & S(t)v(t)\gamma\rho_{x,v} \\ S(t)v(t)\gamma\rho_{x,v} & \gamma^2 v(t) \end{bmatrix}$$

IT IS NOT AFFINE!

Heston Model

Let us define the log transform: $X(t) = \log S(t)$,

$$\begin{cases} \mathrm{d}X(t) &= \left(r - \frac{1}{2}v(t)\right)\mathrm{d}t + \sqrt{v(t)}\mathrm{d}W_x(t), \\ \mathrm{d}v(t) &= \kappa\left(\overline{v} - v(t)\right)\mathrm{d}t + \gamma\sqrt{v(t)}\mathrm{d}W_v(t). \end{cases}$$

Express the model in two independent Brownian motions

$$\left[\begin{array}{c} \mathrm{d}X(t) \\ \mathrm{d}v(t) \end{array}\right] = \left[\begin{array}{c} r - \dfrac{1}{2}v(t) \\ \kappa(\bar{v} - v(t)) \end{array}\right] dt + \left[\begin{array}{cc} \sqrt{v(t)} & 0 \\ \rho_{x,v}\gamma\sqrt{v(t)} & \gamma\sqrt{(1-\rho_{x,v}^2)v(t)} \end{array}\right] \left[\begin{array}{c} \mathrm{d}\widetilde{W}_x(t) \\ \mathrm{d}\widetilde{W}_v(t) \end{array}\right]$$

where Brownian motions $\widetilde{W}_x$ and $\widetilde{W}_v$ are independent.
The instantaneous covariance matrix:

$$\bar{\sigma}(\mathbf{X(t)})\bar{\sigma}(\mathbf{X(t)})^T = \left[\begin{array}{cc} v(t) & v(t)\gamma\rho_{x,v} \\ v(t)\gamma\rho_{x,v} & \gamma^2 v(t) \end{array}\right] \qquad \text{AFFINE!}$$

Characteristic Functions Heston Model

- For Lévy and Heston models, the ChF can be represented by

$$\begin{aligned}\phi(u;\mathbf{x}) &= \varphi_{levy}(u)\cdot \mathrm{e}^{iu\mathbf{x}} \quad \text{with} \quad \varphi_{levy}(u) := \phi(u;0),\\ \phi(u;\mathbf{x},v_0) &= \varphi_{hes}(u;v_0)\cdot \mathrm{e}^{iu\mathbf{x}},\end{aligned}$$

- The characteristic function of the log-asset price for Heston's model:

$$\begin{aligned}\varphi_{hes}(u;v_0) = \; & \exp\left(iur\tau + \frac{v_0}{\gamma^2}\left(\frac{1-\mathrm{e}^{-D\tau}}{1-G\mathrm{e}^{-D\tau}}\right)(\kappa - i\rho\gamma u - D)\right)\cdot \\ & \exp\left(\frac{\kappa\bar{v}}{\gamma^2}\left(\tau(\kappa - i\rho\gamma u - D) - 2\log(\frac{1-G\mathrm{e}^{-D\tau}}{1-G})\right)\right),\end{aligned}$$

with $D = \sqrt{(\kappa - i\rho\gamma u)^2 + (u^2+iu)\gamma^2}$ and $G = \frac{\kappa - i\rho\gamma u - D}{\kappa - i\rho\gamma u + D}$, and $\tau = T - t_0$.