Stochastic Volatility Modelling: A Practitioner’s Approach

Lorenzo Bergomi

Global Markets Quantitative Research

SOCIETE GENERALE
Corporate & Investment Banking
lorenzo.bergomi@sgcib.com

Paris, Third SMAI European School in Financial Mathematics
August 2010
Outline

- Motivation
- Traditional models – the Heston model as an example
- Practitioner’s approach – an example
- Conclusion

Papers *Smile Dynamics I, II, III, IV* are available on SSRN website
Motivation

Why don’t we just delta-hedge options?

Daily P&L of delta-hedged short option position is:

$$P&L = -\frac{1}{2} S^2 \frac{d^2 P}{dS^2} \left( \frac{\delta S^2}{S^2} - \hat{\sigma}^2 \delta t \right)$$

Write daily return as: $$\frac{\delta S_i}{S_i} = \sigma_i Z_i \sqrt{\delta t}$$. Total P&L reads:

$$P&L = -\frac{1}{2} \sum S_i^2 \frac{d^2 P}{dS^2} \left| _i \left( \sigma_i^2 Z_i^2 - \hat{\sigma}^2 \right) \delta t \right.$$  

Variance of daily P&L has two sources:

- the $Z_i$ have thick tails
- the $\sigma_i$ are correlated and volatile
- Delta-hedging not sufficient in practice

Options are hedged with options!
Motivation

The Heston model

Practitioner’s approach – an example

Example 1: barrier option
Example 2: cliquet
Modelling the full volatility surface
Forward variances

Conclusion

Example 1: barrier option
Example 2: cliquet
Modelling the full volatility surface
Forward variances

- Implied volatilities of market-traded options (vanilla, ...) appear in pricing function \( P(t, S, \hat{\sigma}, p, \ldots) \).

▷ Other sources of P&L:

\[
P&L = -\frac{1}{2} S^2 \frac{d^2 P}{dS^2} \left[ \frac{\delta S^2}{S^2} - \hat{\sigma}^2 \delta t \right] - \frac{dP}{d\hat{\sigma}} \delta \hat{\sigma} \\
- \left[ \frac{1}{2} \frac{d^2 P}{d\hat{\sigma}^2} \delta \hat{\sigma}^2 + \frac{d^2 P}{dSd\hat{\sigma}} \delta S \delta \hat{\sigma} \right] + \cdots
\]

- Dynamics of implied parameters generates P&L as well
- Vanilla options should be considered as hedging instruments in their own right

▷ Using options as hedging instruments:

- lowers exposure to dynamics of \textit{realized} parameters, e.g. volatility
- generates exposure to dynamics of \textit{implied} parameters
Example 1: barrier option

In the Black-Scholes model, a barrier option with payoff $f$ can be statically replicated by a European option with payoff $g$ given by:

\[
\text{Barrier: } \begin{cases} 
  f(S) & \text{if } S < L \\
  0 & \text{if } S > L
\end{cases} \quad \text{European payoff: } \begin{cases} 
  f(S) & \text{if } S < L \\
  -\left(\frac{L}{S}\right)^{2r} - 1 & \text{if } S > L
\end{cases}
\]

In our example $f(S) = 1$ and $L = 120$. European payoff is approximately double European Digital.

- Gamma / Vega well hedged by double Euro digital – are there any residual risks?
When $S$ hits 120, unwind double Euro digital. Value of Euro digital depends on implied skew at barrier.

Value of double Euro digital:

$$D = \frac{\text{Put}_{L+\epsilon} - \text{Put}_{L-\epsilon}}{2\epsilon} = \left. \frac{d\text{Put}_K}{dK} \right|_L$$

$$\frac{d\text{Put}_K}{dK} = \frac{d\text{Put}^{BS}_K(K, \hat{\sigma}_K)}{dK} = \frac{d\text{Put}^{BS}_K}{dK} + \frac{d\text{Put}^{BS}_K}{d\hat{\sigma}} \frac{d\hat{\sigma}_K}{dK}$$

$$D = D^{BS}(\hat{\sigma}_L) + \left. \frac{d\text{Put}^{BS}_L}{d\hat{\sigma}} \frac{d\hat{\sigma}_K}{dK} \right|_L$$

Barrier option price depends on scenarios of implied skew at barrier!
Example 2: cliquet

- A cliquet involves ratios of future spot prices – ATM forward option pays:

\[
\left( \frac{S_{T_2}}{S_{T_1}} - k \right)^+
\]

- In Black-Scholes model, price is given by: \( P_{BS}(\hat{\sigma}_{12}, r, \ldots) \)
  - \( S \) does not appear in pricing function ??
  - Cliquet is in fact an option on forward volatility. For ATM cliquet \((k = 100\%)\):

\[
P_{BS} \approx \frac{1}{\sqrt{2\pi}} \hat{\sigma}_{12} \sqrt{T_2 - T_1}
\]

▷ Price of cliquet depends on dynamics of forward implied volatilities
Modelling the full volatility surface

- **Natural approach:** write dynamics for prices of vanilla options as well:

\[
\begin{align*}
    dS &= (r - q) S t + \sigma S dW_t^S \\
    dC^{KT} &= r C^{KT} dt + \bullet \ dW_t^{KT}
\end{align*}
\]

Better: write dynamics on implied vols directly (P. Schönbucher)

\[
\begin{align*}
    dS &= (r - q) S t + \sigma S dW_t^S \\
    d\hat{\sigma}^{KT} &= \star dt + \bullet \ dW_t^{KT}
\end{align*}
\]

- Drift of $\hat{\sigma}^{KT}$ imposed by condition that $C^{KT}$ be a (discounted) martingale
  - How do we ensure no-arb among options of different $K/T$?

- Other approach: model dynamics of local (implied) volatilities (R. Carmona & S. Nadtochiy, M. Schweizer & J. Wissel)
  - Drift of local (implied) vols is non-local & hard to compute

▶ So far inconclusive – try with simpler objects: Var Swap volatilities
Forward variances

- Variance Swaps are liquid on indices – pay at maturity

\[
\frac{1}{T-t} \sum_{t}^{T} \ln \left( \frac{S_{i+1}}{S_{i}} \right)^2 - \hat{\sigma}_{t}^{T}^2
\]

- \(\hat{\sigma}_{t}^{T}\): Var Swap implied vol for maturity \(T\), observed at \(t\)
- If \(S_{t}\) diffusive \(\hat{\sigma}_{t}^{T}\) also implied vol of European payoff \(-2 \ln \left( \frac{S_{T}}{S_{t}} \right)\)

- Long \(T_{2} - t\) VS of maturity \(T_{2}\), short \(T_{1} - t\) VS of maturity \(T_{1}\). Payoff at \(T_{2}\):

\[
\sum_{T_{1}}^{T_{2}} \ln \left( \frac{S_{i+1}}{S_{i}} \right)^2 - \left( (T_{2} - t) \hat{\sigma}_{t}^{T_{2}}^2 - (T_{1} - t) \hat{\sigma}_{t}^{T_{1}}^2 \right) = \sum_{T_{1}}^{T_{2}} \ln \left( \frac{S_{i+1}}{S_{i}} \right)^2 - (T_{2} - T_{1}) V_{t}^{T_{1}T_{2}}
\]

where *discrete* forward variance \(V_{t}^{T_{1}T_{2}}\) is defined as:

\[
V_{t}^{T_{1}T_{2}} = \frac{(T_{2} - t) \hat{\sigma}_{t}^{T_{2}}^2 - (T_{1} - t) \hat{\sigma}_{t}^{T_{1}}^2}{T_{2} - T_{1}}
\]

- Enter position at \(t\), unwind at \(t + \delta t\). P&L at \(T_{2}\) is:

\[
P&L = (T_{2} - T_{1}) \left( V_{t}^{T_{1}T_{2}} - V_{t}^{T_{1}T_{2}} \right)
\]

No \(\delta t\) term in P&L: \(\triangleright V_{t}^{T_{1}T_{2}}\) **has no drift.**
• Replace finite difference by derivative: introduce continuous forward variances $\zeta^T_t$:

$$
\zeta^T_t = \frac{d}{dT} \left((T - t) \hat{\sigma}_t^T \right)^2
$$

$\zeta^T$ is driftless:

$$
d\zeta^T_t = \bullet dW^T_t
$$

• $\zeta^T$ easier to model than $\hat{\sigma}^{KT}$
  - The $\zeta^T$ are driftless
  - Only no-arb condition: $\zeta^T > 0$

▷ Model dynamics of forward variances
Motivation
The Heston model
Practitioner’s approach – an example
Conclusion
Example 1: barrier option
Example 2: cliquet
Modelling the full volatility surface
Forward variances

Full model

- Instantaneous variance is \( \zeta_t^T = t \). Simplest diffusive dynamics for \( S_t \) is:

\[
dS_t = (r - q)S_t dt + \sqrt{\zeta_t} S_t dZ_t^S
\]

- Pricing equation is:

\[
\frac{dP}{dt} + (r - q)S \frac{dP}{dS} + \frac{\zeta_t}{2} S^2 \frac{d^2 P}{dS^2} + \frac{1}{2} \int_t^T \int_t^T \frac{d \zeta_u^t d \zeta_v^t}{dt} \frac{d^2 P}{\delta \zeta_u^t \delta \zeta_v^t} dudv + \int_t^T \frac{dS_t d \zeta_u^t}{dt} \frac{d^2 P}{dS d \zeta_u^t} du = rP
\]

- Dynamics of \( S / \zeta^T \) generates joint dynamics of \( S \) and \( \zeta^{KT} \)

  ▶ Even though VSs may not be liquid, we can use forward variances to drive the dynamics of the full volatility surface.

- Can we come up with non-trivial low-dimensional examples of stochastic volatility models?

- How do we specify a model – what do require from model?
Historical motivations

Traditionally other motivations put forward – not always relevant from practitioner’s point of view – for example:

- **Stoch. vol. needed because realized volatility is stochastic, exhibits clustering, etc.**
  - We don’t care about dynamics of realized vol – we’re hedged. What we need to model is the dynamics of implied vols.

- **Stoch. vol. needed to fit vanilla smile**
  - Not always necessary to fit vanilla smile – usually mismatch can be charged as hedging cost
  - Beware of calibration on vanilla smile:
    - OK if one is able to pinpoint vanillas to be used as hedges.
    - Letting vanilla smile – through model filter – dictate dynamics of implied vols may not be reasonable.
Connection to traditional approach to stochastic volatility modelling

Traditionally stochastic volatility models have been specified using the instantaneous variance:

- Start with historical dynamics of instantaneous variance:
  \[ dV = \mu(t, S, V, p)dt + \alpha()dW_t \]
- In "risk-neutral dynamics", drift of \( V_t \) is altered by "market price of risk":
  \[ dV = (\mu(t, S, V, p) + \lambda(t, S, V))dt + \alpha()dW_t \]
- A few lines down the road, jettison "market price of risk" and conveniently decide that risk-neutral drift has same functional form as historical drift – except parameters now have stars:
  \[ dV = \mu(t, S, V, p^*)dt + \alpha()dW_t \]
- Eventually calibrate (starred) parameters on smile and live happily ever after.

\( V \) is in fact wrong object to focus on – drift issue is pointless:

\[ V_t = \zeta^T_t \rightarrow dV_t = \frac{d\zeta^T_t}{dT} \bigg|_{T=t} dt + \bullet dW_t^t \]
Among traditional models, the Heston model (Heston, 1993) is the most popular:

\[
\begin{align*}
\frac{dV_t}{V_0} &= -k(V_t - V_0)dt + \sigma \sqrt{V_t}dZ_t \\
\frac{dS_t}{S_0} &= (r - q)S_t dt + \sqrt{V_t}S_t dW_t
\end{align*}
\]

- It is an example of a 1-factor Markov-functional model of fwd variances: \(\hat{\sigma}^T\) and \(\zeta^T\) are functions of \(V_t\):

\[
\begin{align*}
\zeta^T_t &= E_t[V_T] = V_0 + (V_t - V_0)e^{-k(T-t)} \\
\hat{\sigma}^T_t &= \frac{1}{T-t} \int_t^T \zeta^T_\tau d\tau = V_0 + (V_t - V_0) \frac{1-e^{-k(T-t)}}{k(T-t)}
\end{align*}
\]

- Look at term-structure of volatilities of \(\hat{\sigma}^T_t\). Dynamics of \(\hat{\sigma}^T_t\) is given by:

\[
\frac{d[\hat{\sigma}^T_t]}{dt} = \ast dt + \frac{1 - e^{-k(T-t)}}{k(T-t)} \sigma \sqrt{V_t}dZ_t
\]
Volatilities of volatilities

- Term-structure of volatilities of volatilities:

\[
T - t \ll \frac{1}{k} \quad \text{Vol}(\sigma_t^T) \approx 1 - \frac{k(T-t)}{2}
\]

\[
T - t \gg \frac{1}{k} \quad \text{Vol}(\sigma_t^T) \approx \frac{1}{k(T-t)}
\]

- Term-structure of historical volatilities of volatilities for the Stoxx50 index:
Term-structure of skew

- ATM skew in Heston model: at order 1 in volatility-of-volatility $\sigma$:

$$
T - t \ll \frac{1}{k} \quad \text{and} \quad \frac{d \hat{\sigma}^{KT}}{d \ln K} \bigg|_{K=F} = \frac{\rho \sigma}{4 \sqrt{V_t}}
$$

$$
T - t \gg \frac{1}{k} \quad \text{and} \quad \frac{d \hat{\sigma}^{KT}}{d \ln K} \bigg|_{K=F} = \frac{\rho \sigma}{2 \sqrt{V_0} \ k(1 - t)}
$$

- Short-term skew is flat, long-term skew decays like $1/(T - t)$

- Market skews of indices display $\sim 1/\sqrt{T - t}$ decay:

![95-105 skew Stoxx50 22/07/10](image1)

![95-105 skew SP500 22/07/10](image2)
ATM skew in Heston model at order 1 in volatility-of-volatility $\sigma$:

$$T - t \ll \frac{1}{k} : \quad \left. \frac{d\hat{\sigma}^{K^T}}{d \ln K} \right|_{K=F} = \frac{\rho \sigma}{4\sqrt{V_t}} \approx \frac{\rho \sigma}{4\hat{\sigma}_{\text{ATM}}}$$

- In Heston model short-term skew is inversely proportional to short-term ATM vol

- Historical behavior for Stoxx50 index: (left-hand axis: $\hat{\sigma}_{\text{ATM}}$, right-hand axis: $\hat{\sigma}_{K=95} - \hat{\sigma}_{K=105}$)

- Maybe not reasonable to hard-wire inverse dependence of skew on $\hat{\sigma}_{\text{ATM}}$. 
In Heston model short ATM vol is normal:

\[ \hat{\sigma}_{ATM} \sim \sqrt{V} \quad \rightarrow \quad d\hat{\sigma}_{ATM} = \star dt + \frac{\sigma}{2} dZ \]

Historical behavior for Stoxx50 index: (left-hand axis: \( \hat{\sigma}_{ATM} \), right-hand axis: 6-month vol of \( \hat{\sigma}_{ATM} \))

\[ \hat{\sigma}_{ATM} \text{ seems log-normal – or more than log-normal – rather than normal.} \]

Other issue: in Heston model VS variances are floored:

\[ \hat{\sigma}_{T}^2 = V_0 + (V_t - V_0) \frac{1 - e^{-k(T-t)}}{k(T-t)} \geq V_0 \frac{k(T-t) - 1 + e^{-k(T-t)}}{k(T-t)} \]
Smile of vol-of-vol – VIX market

- VIX index is published daily: it is equal to the 30-day VS volatility of the S&P500 index: $VIX_t = \hat{\sigma}_t^{t+30 \text{ days}}$
- VIX futures have monthly expiries - their settlement value is the VIX index at expiry

- VIX options have same expiries as futures

\[
F_t^i = E_t[\hat{\sigma}_i^{i+30d}]
\]

\[
C_t^iK = E_t[\left((\hat{\sigma}_i^{i+30d} - K)^+\right]$

VIX futures - 22/07/2010

Smiles of VIX futures - 22/07/2010
From a practitioner’s point of view, question is: what do we require from a model?

Which risks would we like to have a handle on?

- forward skew
- volatilities-of-volatilities, smiles of vols-of-vols
- correlations between spot and implied volatilities
- ...

In next few slides an example of how to proceed to build model that satisfies (some of) our requirements
Practitioner’s approach – an example

- Start with dynamics of fwd variances – we would like a time-homogeneous model
  
  - Start with 1-factor model:
    \[
    d\zeta^T_t = \omega(T - t)\zeta^T_t d\Upsilon_t \rightarrow \ln\left(\frac{\zeta^T_t}{\zeta^T_0}\right) = \bullet + \int_0^t \omega(T - \tau) d\Upsilon_\tau
    \]

  - For general volatility function \(\omega\), curve of \(\zeta^T\) depends on path of \(\Upsilon_t\)

  - Choose exponential form: \(\omega(T - t) = \omega e^{-k(T-t)}\)

  \[
  \int_0^t \omega(T - \tau) d\Upsilon_\tau = \omega e^{-k(T-t)} \int_0^t e^{-k(t-\tau)} d\Upsilon_\tau
  \]

  - Model is now one-dimensional – curve of \(\zeta^T\) is a function of one factor

  - For \(T - t \gg \frac{1}{k}\), at order 1 in \(\omega\):

    \[
    \text{vol}(\sigma^T_t) \propto \frac{1}{k(T-t)} \quad \text{and} \quad \left. \frac{d\sigma^{KT}}{d\ln K} \right|_{K=F} \propto \frac{1}{k(T-t)}
    \]

- No flexibility on term-structure of vols-of-vols and term-structure of ATM skew
Try with 2 factors:

\[ d\zeta_t^T = \omega \zeta_t^T [(1 - \theta) e^{-k_1(T-t)} dW_t^X + \theta e^{-k_2(T-t)} dW_t^Y] \]

- Expression of fwd variances:
  \[ \zeta_t^T = \zeta_0^T e^{\omega x_t^T - \frac{\omega^2}{2} E[x_t^T]} \]
  with \( x_t^T \) given by:
  \[ x_t^T = (1 - \theta) e^{-k_1(T-t)} X_t + \theta e^{-k_2(T-t)} Y_t \]
  \[ dX_t = -k_1 X_t dt + dW_t^X \]
  \[ dY_t = -k_2 Y_t dt + dW_t^Y \]

- Dynamics is low-dimensional Markov – fwd variances are functions of 2 easy-to-simulate factors:
  \[ V_{T1}^{T2} = \frac{1}{T_2 - T_1} \int_{T_1}^{T2} \zeta_t^T dT \]

- Log-normality of \( \zeta^T \) can be relaxed while preserving Markov-functional feature
By suitably choosing parameters, it is possible to mimic power-law behavior for:

- **Term-structure of vol-of-vol**
  - for flat term-structure of VS vols, volatility of VS volatility is given by:
    \[
    \text{vol}(\hat{\sigma}^T)^2 = \frac{\omega^2}{4} \left[ (1 - \theta)^2 \left( \frac{1 - e^{-k_1 T}}{k_1 T} \right)^2 + \theta^2 \left( \frac{1 - e^{-k_2 T}}{k_2 T} \right)^2 
    \right.
    \]
    \[
    \left. + 2\rho_{XY} \theta (1 - \theta) \frac{1 - e^{-k_1 T}}{k_1 T} \frac{1 - e^{-k_2 T}}{k_2 T} \right]
    \]

- **Term-structure of ATM skew**
  - for flat term-structure of VS vols, at order 1 in $\omega$, skew is given by:
    \[
    \left. \frac{d\hat{\sigma}^{KT}}{d \ln K} \right|_F = \frac{\omega}{2} \left[ (1 - \theta)\rho_{SX} \frac{k_1 T - (1 - e^{-k_1 T})}{(k_1 T)^2} + \theta \rho_{SY} \frac{k_2 T - (1 - e^{-k_2 T})}{(k_2 T)^2} \right]
    \]
Term-structure of volatilities of VS vols

- Note that factors have no intrinsic meaning – only vol/vol and spot/vol correlation functions do have physical significance.

It is possible to get slow decay of vol-of-vol and skew.
Conclusion

- Models for exotics need to capture joint dynamics of spot and implied volatilities
- Calibration on vanilla smile not always a criterion for choosing model & model parameters
  - We need to have direct handle on dynamics of volatilities
  - Some parameters cannot be locked with vanillas: need to be able to choose them
- Availability of closed-form formulæ not a criterion either
  - Wrong / unreasonable dynamics too high a price to pay
  - What’s the point in ultrafast mispricing?
- So far, models for the (1-dimensional) set of forward variances. Next challenge: add one more dimension.
- One fundamental issue: in what measure does the initial configuration of asset prices – e.g. implied volatilities – restrict their dynamics?