Smile dynamics: modeling the smile of volatility (of volatility)

Lorenzo Bergomi

Equity Derivatives Quantitative Research
Société Générale
lorenzo.bergomi@sgcib.com

Roma, Global Derivatives 2009
Outline

- Rationale for new type of model
- A simple framework for the dynamics of fwd variances
- Generating smiles for fwd variances
- Using the model
  - Calibration to VIX smiles
  - Application to swaptions & options on realized variance
  - The vanilla skew
- The Skew Stickiness Ratio
- Conclusion
Motivation - 1

- New listed volatility instruments
  - VIX futures & options

- New OTC options on variance
  - Swaptions
  - Options on realized variance, spot or forward-starting
  - Multi-asset options on realized variance
  - Hybrid equity / realized variance payoffs

⇒ Require modeling of
  - joint dynamics of spot / variance curve
  - smile of forward variances

\[
S = V_{T_1}^{T_2}, \quad S = \frac{1}{T_2 - T_1} \sum_{i=T_1}^{T_2} \ln \left( \frac{S_{i+1}}{S_i} \right)^2
\]
Motivation - 2

- 1-factor stoch. vol models generate poor smile dynamics
  - One single time scale: for $T >> 1/k$, $\sigma_{Vol} \propto \frac{1}{T}$, $\text{Skew}_{ATM} \propto \frac{1}{T}$

- Popular smile models have other structural problems
  - Heston model:
    - short ATM vol is normal
    - short ATM skew inv. prop. to level of ATM vol
    - simulation of dynamics of variance curve uselessly arduous
  - Jump / Lévy models:
    - forward skew but no dynamics of vol / skew
    - process for producing skew generates unrealistic skewness for short-term returns
      $\Rightarrow$ pbs for very path-dependent options
  - Stoch. Vol Lévy
    - dynamics is very constrained: for short maturities $\frac{d\hat{\sigma}}{d\ln K} \bigg|_F \propto \frac{1}{\hat{\sigma}_F}$
A simple framework - 1

- Start with simple lognormal dynamics for forward variances driven by OU factors:

\[
\xi_t^T = \xi_0^T e^{\omega x_t^T - \frac{\omega^2}{2} E[(x_t^T)^2]} \quad \xi_0^T = \frac{d}{dT} \left( T \hat{\sigma}^2_{t=0}(T) \right)
\]

\[
x_t^T = (1 - \theta) e^{-k_1(T-t)} X_t + \theta e^{-k_2(T-t)} Y_t
\]

\[
dX = -k_1 X dt + dW^X
\]

\[
dY = -k_2 Y dt + dW^Y
\]

\[X_t, Y_t\] are Gaussian: correlation of \( W^X \) and \( W^Y \) is \( \rho_{XY} \)

- Dynamics is Markovian:

\[
\frac{dV_t^{T_1, T_2}}{V_t^{T_1, T_2}} = \alpha_{X}^{T_1, T_2} dW^X + \alpha_{Y}^{T_1, T_2} dW^Y
\]

\[
\alpha_{X}^{T_1, T_2} = (1 - \theta) \frac{\int_{T_1}^{T_2} \xi_t^T e^{-k_1(T-t)} dT}{\int_{T_1}^{T_2} \xi_t^T dT} , \quad \alpha_{Y}^{T_1, T_2} = \theta \frac{\int_{T_1}^{T_2} \xi_t^T e^{-k_2(T-t)} dT}{\int_{T_1}^{T_2} \xi_t^T dT}
\]
A simple framework - 2

- **We can use same dynamics for discrete forward variances**
  - Allows for independent control of short fwd skew

- **2-factor model is minimal setup that provides handle on:**
  - term-structure of volatilities of volatilities
  - term-structure of skew

- Adding more factors generates no additional complexity:

$$
\xi_t^T = \xi_0^T e^{\omega x_t^T} - \frac{1}{2} \omega^2 E[(x_t^T)^2]
$$

$$
\chi_t^T = \sum_{n} w_n e^{-k_n(T-t)} \chi_n^T
$$
Smile for fwd variances – continuous 1

- Consider instantaneous fwd variances $\xi^T$
- Would like to keep model Markovian while relaxing lognormality:
  
  - Define
  \[ x_t^T = \alpha_\theta \left( (1 - \theta) e^{-k_1 (T-t)} X_t + \theta e^{-k_2 (T-t)} Y_t \right) \]
  
  with \( \alpha_\theta = \frac{1}{\sqrt{(1 - \theta)^2 + \theta^2 + 2 \rho_{XY} \theta (1 - \theta)}} \)

  - Write
  \[ \xi_t^T = \xi_0^T f^T (x_t^T, t) \]

  - Conditions on \( f^T \):
    - \( \xi_t^T \) is driftless:
    \[ \frac{df^T}{dt} + \frac{\lambda^2 (T-t)}{2} \frac{d^2 f^T}{dx^2} = 0 \]
    
    with
    \[ \lambda^2 (\tau) = \alpha_\theta^2 \left[ (1 - \theta)^2 e^{-2k_1 \tau} + \theta^2 e^{-2k_2 \tau} + 2 \rho_{XY} \theta (1 - \theta) e^{-(k_1 + k_2) \tau} \right] \]
    
    - \( f^T (0, 0) = 1 \)
    - \( f^T \) monotonic in \( x \)

  - Ex: lognormal dynamics given by
  \[ f^T (x, t) = e^{\omega x - \frac{\omega^2}{2} \int_{T-t}^T \lambda^2 (\tau) d\tau} \]
Smile for fwd variances – continuous 2

- Analogous to "Markov-functional" models of fixed income
  - Is in fact a local vol model:
    \[ \sigma(\zeta^T, t) = \eta(T-t) \frac{\partial \ln f^T}{\partial x} \bigg|_{x = f^{-1}(\zeta, t)} \]
  - \( f^T \) uniquely determined by terminal profile: \( f^T(x, t = T) \)

- No market smiles for individual \( \zeta^T \), rather for discrete fwd variances: \( V^{T_i}_{T_{i+1}} = \frac{1}{T_{i+1} - T_i} \int_{T_i}^{T_{i+1}} \zeta^T dT \)
  - (1) Assume \( f^T(x, T_i) \) are identical for \( T \in [T_i, T_{i+1}] \): exact calibration on VIX smiles
  - (2) Use analytical ansatz for \( f^T \): representation on basis of space-time harmonic functions
    \[ f^T(x, t) = \int_0^{+\infty} d\mu(\omega) e^{\omega x - \frac{\omega^2}{2} \int_{T-t}^{T} \lambda(\tau)^2 d\tau} \]

Using just 2 exponentials:

\[ f^T(x, t) = (1 - \gamma_T)e^{\omega_T x - \frac{\omega_T^2}{2} \int_{T-t}^{T} \lambda(\tau)^2 d\tau} + \gamma_T e^{(\beta_T \omega_T)x - \frac{(\beta_T \omega_T)^2}{2} \int_{T-t}^{T} \lambda(\tau)^2 d\tau} \]

with \( \omega_T = (2\nu \xi^T) \frac{1}{1 - \gamma_T + \gamma_T \beta_T} \)

Initially vol of very short vol is \( \nu \):
\[ d\xi^T_{t=T} = (2\nu \xi^T) \xi^T dx^T \]
Smile for fwd variances – continuous 3

- Process for spot
  \[ dS = (r - q)Sdt + \sqrt{\xi_t^t} \sqrt{f^t(x^t_t, t)}S dW_t \]
  \[ x^t_t = \alpha_{\theta}(\theta X_t + \theta Y_t) \]

- Vanilla skew is set by correlations \( \rho_{SX}, \rho_{SY} \)

- Pricing is painless
  - Processes \( X_t, Y_t \) are simulated (1) exactly, (2) with no time-stepping
  - Brownian motion for spot process

- Before we turn to examples: discrete version
Smile for fwd variances – discrete 1

- Assume a tenor structure of expiries, for example those of VIX futures & options

- Model discrete fwd variances $V_{T_iT_{i+1}}$:
  \[ V_{T_iT_{i+1}}^{T_{i+1}} = V_0^{T_{i+1}} f^i (x^i_t, t) \]
  \[ x^i_t = \alpha_\theta \left( (1 - \theta) e^{-k_1(T_i-t)} X_t + \theta e^{-k_2(T_i-t)} Y_t \right) \]
  \[ \frac{df^i}{dt} + \frac{\lambda^2 (T_i - t)}{2} \frac{d^2 f^i}{dx^2} = 0 \]

- Determine terminal profile $f^i(x_t = T_i)$ so as to exactly recover market smile for $V_{T_iT_{i+1}}^{T_{i+1}}$
  - Use procedure outlined by J. Kennedy, P. Hunt, A. Pelsser
  - VIX smiles provide market prices for digitals on $\sqrt{V_{T_iT_{i+1}}^{T_{i+1}}}$ hence on $V_{T_iT_{i+1}}^{T_{i+1}}$

\[ D_L = P(V_{T_i}^{T_{i+1}} < L) = P(x^i_T < x(L)) \quad \text{with} \quad f^i(x(L), T_i) = \frac{L}{V_0^{T_{i+1}}} \]
Smile for fwd variances – discrete 2

- **Process for spot over interval** $[T_i, T_{i+1}]$

  - Solution (1): variance over interval $[T_i, T_{i+1}]$ stays constant, equal to $V_{T_i}^{T_{i+1}} = V_0^{T_i} f^i(x_{T_i}^i, T)$

    ⇒ both vanilla skew and fwd skew set by correlations $\rho_{SX}, \rho_{SY}$

  - Solution (2): control fwd skew by choosing "local vol" $\sigma_i(S) \equiv \frac{\phi}{S_{i}^{T_i}} V_{i}^{T_{i+1}}$ such that:
    - VS variance for maturity $T_{i+1}$ is $V_{T_i}^{T_{i+1}}$
    - Desired level of fwd skew over $[T_i, T_{i+1}]$ is obtained
    - Desired dependence of level fwd skew on level of vol is achieved

  - For example, use CEV form: $\varphi(x) = \sigma_0 x^{1-\beta}$

    $\hat{\sigma}_{95\%} - \hat{\sigma}_{105\%} = 5\%$

    $\frac{\hat{\sigma}_{95\%} - \hat{\sigma}_{100\%}}{\hat{\sigma}_{100\%}} = 0.25$

[Graphs and equations related to skew and volatility parameters]
Smile for fwd variances – discrete 3

- Benefits of using discrete version:
  - Control on smile of fwd vols – VIX smiles easily calibrated
  - Control on short fwd skew and its dependence on level of fwd vol
  - Correlations $\rho_{X,Y}$ still at our disposal to control vanilla smile independently

- Again pricing is painless:
  - Processes $X_t, Y_t$ are simulated (1) exactly, (2) with no time-stepping
  - Brownian motion for spot process
  - Mapping functions $f'(x, t)$ are generated once and for all
Using the model – calibration to VIX smiles - 1

- Here we use continuous version
- Calibration on VIX smiles:
  - Choose values for \( v, \theta, k_1, k_2, \rho_{XY} \): sets general dynamics of fwd vols in model
  - Use piecewise constant parameters \( \gamma_i, \beta_i, \zeta_i \) on each interval \([T_i, T_{i+1}][\)
    - \( \gamma, \beta \) control skew / \( \zeta \) controls vol level – of fwd vol
  - Use VS term-structure computed on SP500 smile as input
  - Calibrate both
    - VIX futures themselves
      \[ F = E[\sqrt{V_{T_i T_{i+1}}_{T_i}}] \]
    - VIX options
      \[ C_K = E[(\sqrt{V_{T_i T_{i+1}}_{T_i}} - K)^+] \]
  - Is VIX market consistent with SP500 option’s market?
    - 1-month fwd variance can be replicated with VIX futures & options
      \[ V_{T_i T_{i+1}} = F^2 + 2\int_0^F P_K dK + 2\int_F^\infty C_K dK \]
Using the model – calibration to VIX smiles – 2

- Pricing of options on fwd vol/var easy: simple Gaussian integration

**March 18, 2008**

![Graph showing market and model smiles comparison for March 18, 2008.]

<table>
<thead>
<tr>
<th>Date</th>
<th>Market</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>16-Apr-08</td>
<td></td>
<td></td>
</tr>
<tr>
<td>21-May-08</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18-Jun-08</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16-Jul-08</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20-Aug-08</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\gamma$</td>
<td>130.0%</td>
</tr>
<tr>
<td>$\beta$</td>
<td>28%</td>
</tr>
<tr>
<td>$\zeta$</td>
<td>8.0</td>
</tr>
<tr>
<td>$k_1$</td>
<td>0.35</td>
</tr>
<tr>
<td>$k_2$</td>
<td>0%</td>
</tr>
</tbody>
</table>

**Sept 12, 2008**

![Graph showing market and model smiles comparison for Sept 12, 2008.]

<table>
<thead>
<tr>
<th>Date</th>
<th>Market</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>22-Oct-08</td>
<td></td>
<td></td>
</tr>
<tr>
<td>19-Nov-08</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17-Dec-08</td>
<td></td>
<td></td>
</tr>
<tr>
<td>21-Jan-09</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18-Feb-09</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\gamma$</td>
<td>35%</td>
</tr>
<tr>
<td>$\beta$</td>
<td>16%</td>
</tr>
<tr>
<td>$\zeta$</td>
<td>87%</td>
</tr>
</tbody>
</table>

[Graphs and data tables are shown with market and model comparisons for various dates, highlighting the calibration process.]
Using the model – calibration to VIX smiles – 3

• Sept 25, 2008

- Feb 25, 2009

- March 13, 2009
Using the model – options on realized variance – swaptions

• Option on realized variance pays \( \left( \frac{1}{T} \sum_{t=0}^{T} \ln \left( \frac{S_{t+1}}{S_t} \right)^2 - K \right)^+ \), swaption pays \( (V_{T_2}^{T_1} - K)^+ \)

• Popular approximation:
  - Define underlying effective underlying \( Y \):
    \[ Y = \frac{t \sigma_{i}^2 + (T - t) \hat{\sigma}_{i, T-t}^2}{T} \]
  - \( Y \) hedged by trading VS of maturity \( T \): \( Y \) is driftless
  - Dynamics for \( Y \) depends on dynamics of \( \hat{\sigma}_{i, T-t} \)

\( \Rightarrow \) Price of option on variance only depends on dynamics of VS vol of residual maturity.

• Price option with model

\[ \begin{array}{c|c|c|c}
\text{Set1} & \text{Set 2} & \text{Set 3} \\
\hline
v & 130.0\% & 137.0\% & 125.0\% \\
\theta & 28\% & 29\% & 32\% \\
k_1 & 8.0 & 12.0 & 4.5 \\
k_2 & 0.35 & 0.30 & 0.60 \\
\rho_{XY} & 0\% & 90\% & -70\% \\
\end{array} \]

<table>
<thead>
<tr>
<th>Spot-starting realised</th>
<th>Set1</th>
<th>Set 2</th>
<th>Set 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>6 months</td>
<td>2.03%</td>
<td>2.03%</td>
<td>2.03%</td>
</tr>
<tr>
<td>1 year</td>
<td>2.22%</td>
<td>2.21%</td>
<td>2.24%</td>
</tr>
<tr>
<td>6 months realized in 6 months</td>
<td>3.04%</td>
<td>2.89%</td>
<td>3.25%</td>
</tr>
<tr>
<td>6 months in 6 months swaption</td>
<td>2.28%</td>
<td>2.10%</td>
<td>2.57%</td>
</tr>
</tbody>
</table>

\( \Rightarrow \) Pop. approx ok for spot-starting options on realized variance
Using the model – dynamics of fwd vols

<table>
<thead>
<tr>
<th></th>
<th>Set 1</th>
<th>Set 2</th>
<th>Set 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\nu$</td>
<td>130.0%</td>
<td>137.0%</td>
<td>125.0%</td>
</tr>
<tr>
<td>$\theta$</td>
<td>28%</td>
<td>29%</td>
<td>32%</td>
</tr>
<tr>
<td>$k_1$</td>
<td>8.0</td>
<td>12.0</td>
<td>4.5</td>
</tr>
<tr>
<td>$k_2$</td>
<td>0.35</td>
<td>0.30</td>
<td>0.60</td>
</tr>
<tr>
<td>$\rho_{XY}$</td>
<td>0%</td>
<td>90%</td>
<td>-70%</td>
</tr>
</tbody>
</table>

⇒ term-structure of vols of spot-starting vols does not uniquely determine vols & correls of fwd vols

Spot-starting variance = basket of fwd variances. Same basket vol can be achieved with:

- low vols / high correls
- high vols / low correls

For flat term-structure of VS vols, inst. vol of VS vol given by:

$$\sigma_T^{vol} = v\alpha_\theta \sqrt{\theta^2 \left(1 - e^{-k_1T}/k_1\right)^2 + (1 - \theta)^2 \left(1 - e^{-k_2T}/k_2\right)^2 + 2\rho_{XY}\theta(1 - \theta) \left(1 - e^{-k_1T}/k_1\right)\left(1 - e^{-k_2T}/k_2\right)}$$

Societe Generale
Corporate & Investment Banking
Using the model – options on realized variance – smile

- Use parameters calibrated on VIX smiles of March 18, 2008

- Price option on realized variance – back out smile:
Using the model – the vanilla skew – 1

• Vanilla skew in continuous version is generated by spot/vol correlations $\rho_{SX}, \rho_{SY}$

• Approximate expression for vanilla skew:
  - ATM skew is related to skewness $\Sigma_{T}$ of $x = \ln \left( \frac{S_T}{S_0} \right)$ through:
    \[
    \left. \frac{d\hat{\sigma}}{d \ln K} \right|_F = \frac{\Sigma_{T}}{6\sqrt{T}}
    \]
  - is given at 1st order in vol-of-vol by (double) integral of spot/vol correlation function:
    \[
    \left< (x-<x>)^3 \right> = 6\sigma_0 \int_0^T dt \int_0^t \left< \frac{dS_t}{S_t}, \sigma_t \right>
    \]
  \Rightarrow yields expression of ATMF skew at order 1 in vol of vol:
    \[
    \left. \frac{d\hat{\sigma}}{d \ln K} \right|_F = \nu \alpha \left[ (1 - \theta)\rho_{SX} \frac{k_1 T - (1 - e^{-k_1 T})}{k_1^2 T^2} + \theta \rho_{SY} \frac{k_2 T - (1 - e^{-k_2 T})}{k_2^2 T^2} \right]
    \]

• Example: 95% - 105% skew with $\rho_{SX} = -70\%, \rho_{SY} = -36\%$

\Rightarrow Parameters generating slow decay of vol of vol also generate slow decay of skew

95%-105% skew

- Actual
- Approximate

<table>
<thead>
<tr>
<th>$\rho_{SX}$</th>
<th>-70%</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho_{SY}$</td>
<td>-36%</td>
</tr>
</tbody>
</table>

\[
\rho_{SX} = -70\%, \rho_{SY} = -36\%
\]

Skew $\propto \frac{1}{\sqrt{T}}$
The Skew Stickiness Ratio – 1

• Natural question: how does the VS or ATMF vol move when spot moves?
  - in units of the ATMF skew?

=> enter the Skew Stickiness Ratio

\[ R_T = \frac{\frac{\delta \sigma_{VS}^T}{\delta \ln S}}{d \hat{\sigma}} \bigg|_{F,T} \]

with
\[ \frac{\delta \sigma_{VS}}{\delta \ln S} = \frac{\langle \delta \sigma_{VS} \delta \ln S \rangle}{\langle (\delta \ln S)^2 \rangle} \]

• At 1st order in vol of vol, with flat VS term-structure:
  • For short maturities \( R_T \rightarrow 2 \) : model-independent
  • For (very) long maturities \( R_T \rightarrow 1 \)

• General expression for any number of factors:

\[ R_T = \sum w_i \rho_{iS} \frac{1 - e^{-k_i T}}{k_i T} \]

\[ = \frac{\sum w_i \rho_{iS} k_i T - (1 - e^{-k_i T})}{(k_i T)^2} \]

\[ \text{with } \rho_{iS} = 130.0\% \]

\[ \rho_Y = 28\% \]

\[ k_1 = 8.0 \]

\[ k_2 = 0.35 \]

\[ \rho_{XY} = 0\% \]

\[ \rho_{SY} = -70\% , \rho_{SV} = -36\% \]

Maturity (years)

Lorenzo Bergomi
The Skew Stickiness Ratio – 2

• Model-independent (no jumps) expression of the SSR at order 1 in the vol-of-vol (flat term-structure of vols):

\[ R_T = \frac{1}{T} \int_0^T f(t)dt \]

\[ \frac{1}{T^2} \int_0^T dt \int_0^t f(\tau)d\tau \]

where \( f(\tau) \) is the spot/vol correlation function:

\[ f(\tau) = \frac{1}{dt} \left( \frac{dS_t}{S_t} \sigma_{t+\tau} \right) \]

• If \( f \) is monotonic then \( 1 \leq R \leq 2 \)

• For \( T \to 0 \) above expression recovers well-known model-independent result that \( R \to 2 \)
The Skew Stickiness Ratio – 3

• Scaling of \( f(\tau) \) for large \( \tau \) governs
  ➞ The speed with which the ATMF skew decays
  ➞ The value of the SSR for large \( T \)

• Imagine that for large \( t \) \( f(\tau) \propto \frac{1}{\tau^\beta} \) then :

  - If \( \beta > 1 \) \( \frac{d\hat{\sigma}}{d\ln K}\bigg|_F \propto \frac{1}{T} \) and \( R_T \to 1 \)

  - If \( \beta < 1 \) \( \frac{d\hat{\sigma}}{d\ln K}\bigg|_F \propto \frac{1}{T^\beta} \) and \( R_T \to 2 - \beta \)

  Exponential decay same as \( \beta > 1 \)
  ➞ The decay of the ATMF skew (static feature) and the large-\( T \) value of the SSR (dynamic feature) are related for large \( T \) through:

  \[
  \text{Skew}_T \propto \frac{1}{T^{2-R}}
  \]

  ➞ Explains why in our example:

  • intermediate regime where \( \text{Skew}_T \propto 1/T^{0.5}, R_T \approx 1.5 \)
  • eventually for long maturities \( \text{Skew}_T \propto 1/T, R_T \approx 1 \)
Conclusion

• 2-factor stoch vol model allows control over dynamics of fwd vols
• Model is driven by easy-to-simulate Gaussian processes – no sweating
• Can be extended to $N$ factors with no additional complexity
  - needed for options on slope of VS curve
• Allows control on smile of fwd vols – almost for free
  - can be calibrated to VIX smiles, exactly in discrete version
• Discrete version allows additional control on short fwd skew and decoupling from vanilla skew
  - magnitude & dependence on level of short fwd vol
• Parameters governing dynamics of fwd vols also govern decay of vanilla skew
• Model-independent relationship at 1st order in the vol-of-vol links SSR to decay of the ATMF skew