Normal Expansion of SABR

Philippe Balland, UBS

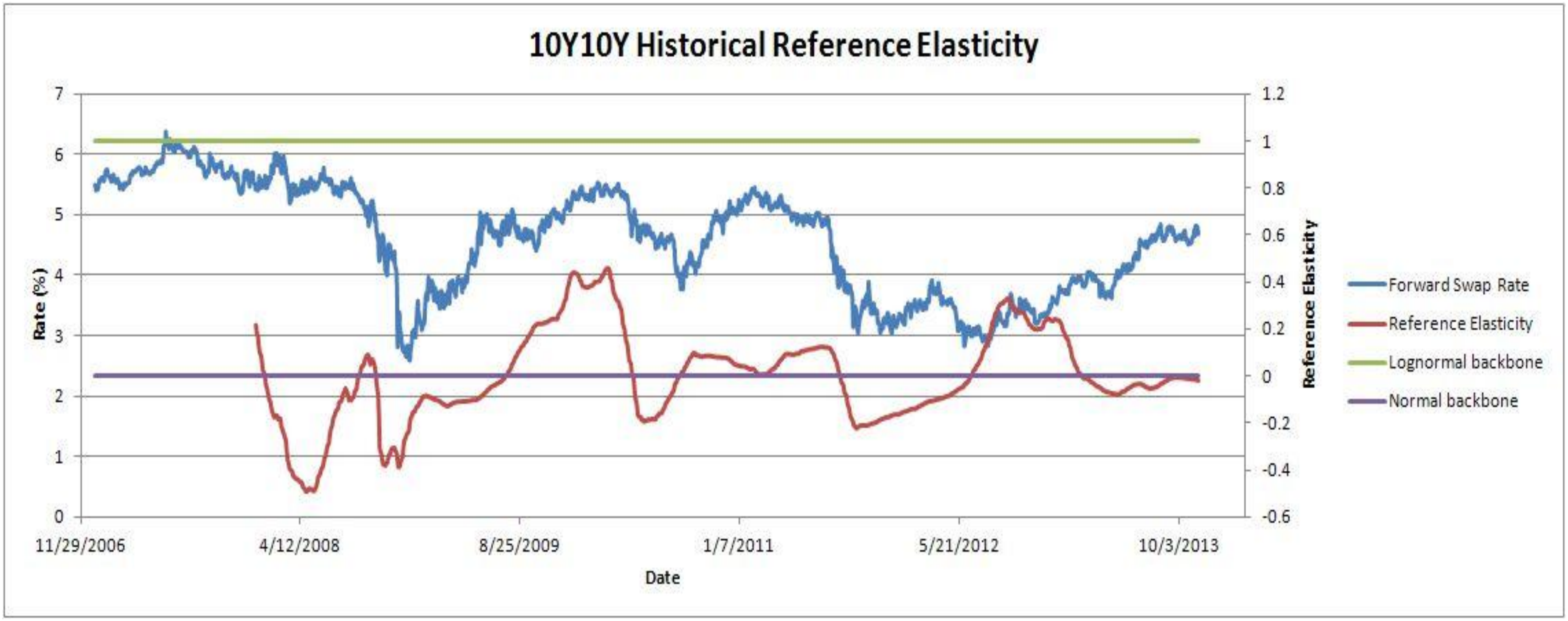
Philippe.Balland@ubs.com
ppg.balland:gmail.com

Introduction

Historical and implied backbones define the 'expected' relationship between swap rates and ATM volatilities:

$$\beta_{ref} = 1 + \frac{\Delta \ln ATM}{\Delta \ln S}.$$

Typical interest rates have a backbone closer to normal than lognormal:



The SABR formula is based on an asymptotic expansion of the *Lognormal* implied volatility. However, this formula becomes inaccurate and can imply negative density when the elasticity parameter is small. We propose to resolve this issue by performing an asymptotic expansion of the *Normal SABR* implied volatility instead.

Local Volatility for Normal SABR

Consider a forward swap rate *S* following the Normal SABR dynamic:

$$dS_t = b_0 \alpha_t dW_{t\wedge\tau},$$

$$\alpha_t = \exp\left(\gamma B_t - \frac{1}{2}\gamma^2 t\right),$$

where τ is the first time the swap rate hits $S_{min} = 0$.

Normal SABR generates the same smiles as a Quadratic Local Variance model:

$$dS_t = b_0 \sqrt{q(S_0 - S_t)}\, dW_{t\wedge\tau},$$

$$q(U) = 1 - 2\rho\tilde{\gamma}U + \tilde{\gamma}^2 U^2, \quad \tilde{\gamma} = \frac{\gamma}{b_0}.$$

Local Volatility for Normal SABR: Proof

The local volatility is calculated as follows:

$$\sigma(t,K)^2 = 1\{K > S_{min}\}\, b_0^2 \frac{E[\alpha_t^2 \delta(S_t - K)]}{E[\delta(S_t - K)]} \equiv 1\{K > S_{min}\}\, b_0^2 \frac{L}{D}.$$

➢ Calculation of *L*:

Define:

$$U_t = \frac{S_t - K}{\alpha_t},\ \frac{dQ^\alpha}{dQ} = \alpha.$$

Observe:

$$L = E[\alpha_t^2 \delta(S_t - K)] = E\left[\alpha_t \delta\left(\frac{S_t - K}{\alpha_t}\right)\right] = E^\alpha[\delta(U_t)].$$

By Ito's lemma, derive:

$$dU_t = b_0 q(U_t)^{\frac{1}{2}} dW^\alpha_{t\wedge\tau},$$

$$q(U) = 1 - 2\rho\hat{\gamma}U + \hat{\gamma}^2 U^2.$$

➢ Calculation of *D*:

Define:
$$\eta_t = \frac{e^{-\gamma^2 t}}{\alpha_t}, \ \frac{dQ^\eta}{dQ} = \eta.$$

Similarly to above:
$$D = E[\delta(S_t - K)] = e^{\gamma^2 t} E^\eta[\delta(U_t)],$$

$$dU_t = b_0 q(U_t)^{\frac{1}{2}} dW^\eta_{t\wedge\tau} + 1\{t < \tau\} b_0^2 \dot{q}(U_t) dt.$$

Define $\rho_t = q(U_0)/q(U_t) e^{\gamma^2 t}$ and derive:

$$\frac{d\rho_t}{\rho_t} = -b_0 \frac{\dot{q}(U_t)}{q(U_t)^{\frac{1}{2}}} dW^\eta_{t\wedge\tau}.$$

Consider the probability measure Q^ϱ such that $\frac{dQ^\varrho}{dQ^\eta} = \rho$:

$$dU_t = b_0 q(U_t)^{\frac{1}{2}} dW^\rho_{t\wedge\tau}.$$

Obtain the following expression for *D*:

$$D = \frac{e^{\gamma^2 t}}{q(U_0)} E^\eta\left[\frac{q(U_0)}{q(U_t)} \delta(U_t)\right] = \frac{E^\rho[\delta(U_t)]}{q(U_0)} = \frac{L}{q(U_0)}.$$

Derive by grouping the terms for *D* and *L*:

$$\sigma(t, K) = 1\{K > S_{min}\} b_0 \sqrt{q(S_0 - K)} \quad \boxdot$$

Local Volatility for SABR

Consider a forward swap rate *S* following the SABR dynamic:

$$dS_t = \sigma_0 \alpha_t S_t^{\beta} dW_{t\wedge\tau},$$

$$\alpha_t = \exp\left(\gamma B_t - \tfrac{1}{2}\gamma^2 t\right).$$

SABR and the below Local Volatility model imply similar smiles:

$$dS_t = \sigma_0 \sqrt{p\left(d_\beta(S_t, S_0)\right)}\, dW_{t\wedge\tau},$$

$$d_\beta(S, S_0) \equiv \int_{S_0}^{S} \frac{du}{u^\beta} = \frac{S^{1-\beta} - S_0^{1-\beta}}{1-\beta},$$

$$p(U) = 1 + 2\tilde{\gamma}\rho U + \tilde{\gamma}^2 U^2, \quad \tilde{\gamma} = \frac{\gamma}{\sigma_0}.$$

This result is obtained using the same calculation as the one explained above for the Local Volatility for Normal SABR. Details can also be found in *P. Balland and Q. Tran, 'SABR goes normal', Risk, June 2013*.

Normal Expansion of SABR

From the local volatility, derive the asymptotic expansion of the *lognormal* implied volatility:

$$IV_{asymptotic}(K;\beta,\gamma,\rho,\sigma_0) = \sigma_0 \frac{\ln \frac{K}{S_0}}{\int_0^{d_\beta(K,S_0)} \frac{du}{\sqrt{p(u)}}}.$$

Because the Lognormal model has no absorption at zero, this expansion loses accuracy for low β and large expiry, eventually leading to negative density.

A solution is to perform an asymptotic expansion of the *Normal SABR* implied volatility b_0.

By matching asymptotic expansions $IV_{asymptotic}(K;0,\gamma,\rho,b_0) = IV_{asymptotic}(K;\beta,\gamma,\rho,\sigma_0)$, derive:

$$\int_0^{\frac{K-S_0}{b_0}} \frac{du}{\sqrt{1+2\gamma\rho u+\gamma^2 u^2}} = \int_0^{\frac{d_\beta(K,S_0)}{\sigma_0}} \frac{du}{\sqrt{1+2\gamma\rho u+\gamma^2 u^2}}.$$

Therefore, the Normal SABR implied volatility is:

$$b_0 = \sigma_0 \frac{(1-\beta)(K-S_0)}{K^{1-\beta}-S_0^{1-\beta}}.$$

Pricing Formula for Normal SABR

- **Replace Normal SABR by the Quadratic Local Variance process:**

$$dX_t = b_0 p(X_t)^{\frac{1}{2}} dW_{t\wedge\tau} ,$$

$$X_t = S_t - S_0 ,$$

$$p(X) = 1 + 2\rho\tilde{\gamma}X + \tilde{\gamma}^2 X^2, \quad \tilde{\gamma} = \frac{\gamma}{b_0},$$

where τ is the first time the process X_t hits the boundary $S_{min} - S_0$.

- **Apply the Tanaka-Meyer formula:**

$$E[(S_T - K)^+] = (S_0 - K)^+ + \frac{1}{2}\int_0^T b_0^2\, p(x) E[\delta(X_t - x)] dt,$$

$$x = K - S_0.$$

Define:

$$I(X) = \int_0^X \frac{1}{\sqrt{p(u)}} du = \frac{1}{\tilde{\gamma}} \ln\left(\frac{\tilde{\gamma}X + \rho + \sqrt{p(X)}}{1+\rho}\right).$$

Calculate:

$$p(x)E[\delta(X_t - x)] = p(x)^{\frac{1}{2}} E[\delta(I_t - I(x))],$$

$$dI_t = b_0 dW_{t\wedge\tau} - \tfrac{1}{4} 1\{\tau > t\} b_0^2 \frac{\dot{p}(X_t)}{p(X_t)^{\frac{1}{2}}} dt.$$

Resolve the drift by changing measure to Q^θ with $\frac{d\theta_t}{\theta_t} = \frac{1}{4}\frac{\dot{p}(X_t)}{p(X_t)^{\frac{1}{2}}} b_0 dW_{t\wedge\tau}$.

Derive: $$\mathrm{dI_t = b_0 dW^\theta_{t\wedge\tau}}.$$

Observe that $A_t = p(X_t)^{\frac{1}{4}}$ satisfies $d\ln A_t = d\ln\theta_t + 1\{\tau > t\}\left(-\frac{1}{8} + \frac{3}{8}\times\frac{1-\rho^2}{p(X_t)}\right)\gamma^2 dt$.

Derive:

$$p(x)^{\frac{1}{2}} E[\delta(I_t - I(x))] = p(x)^{\frac{1}{4}} E^\theta\left[\exp\left(-\frac{1}{8}\gamma^2 t\wedge\tau + \frac{3}{8}\gamma^2(1-\rho^2)\int_0^{t\wedge\tau}\frac{1}{p(X_u)}du\right)\delta(I_t - I(x))\right].$$

- **Finally, obtain:**

$$[(S_T - K)^+] = (S_0 - K)^+ + \frac{p(K - S_0)^{\frac{1}{4}}}{\sqrt{2\pi}} b_0 \times \frac{1}{2}\int_0^T e^{-\frac{1}{8}\gamma^2 t}\Phi(t, B)\times[e^{\frac{-B^2}{2t}} - e^{\frac{-C^2}{2t}}]dt,$$

$$B = \frac{I(K - S_0)}{b_0}, \quad C = \frac{2I(S_{min} - S_0) - I(K - S_0)}{b_0},$$

$$\Phi(t, z) = E\left[\exp\left(\frac{3}{8}\gamma^2(1-\rho^2)\int_0^t \frac{1}{\varphi(W_u)}du\right)\middle| \mathrm{W_t = z\ \&\ \tau_C} > t\right],$$

$$\varphi(W) = \frac{1}{4}\left((1+\rho)^2 e^{2\gamma W} + (1-\rho)^2 e^{-2\gamma W} + 2(1-\rho^2)\right).$$

Numerical Implementation

➢ **Choose a grid $\{T_i : i = 0, \cdots, N\}$ such that $T_0 = 0,\ T_N = T$:**

$$E[(S_T - K)^+] = (S_0 - K)^+ + \frac{p(K - S_0)^{\frac{1}{4}}}{\sqrt{2\pi}} b_0 \sum_{i=1}^{N} e^{-\left(\frac{1}{8}\gamma^2 + \kappa_i\right) T_{i-1}} \Phi(T_{i-1}, B) \times J_i,$$

$$\kappa_i = -\frac{1}{8}\gamma^2 + \frac{\ln \Phi(T_i, B) - \ln \Phi(T_{i-1}, B)}{T_i - T_{i-1}},$$

$$J_i = \frac{1}{2}\int_{T_{i-1}}^{T_i} \frac{1}{\sqrt{t}} e^{\kappa_i t}\left(e^{\frac{-B^2}{2t}} - e^{\frac{-C^2}{2t}}\right) dt.$$

The integrals J_i can be analytically calculated at a cost similar to two cumulative normal calculations using the formula 7.4.33 in *M. Abramowitz and I.A. Stegun, Handbook of Mathematical Functions, 1972*.

➢ **Simplify the calculation of $\Phi(T_i, B)$ with the following approximation:**

$$\Phi_i(z) \approx E\left[\exp\left(\frac{3}{8}\gamma^2(1-\rho^2)\int_0^{T_i} \frac{1}{\varphi(W_u)} du\right) \middle| \mathrm{W}_{\mathrm{T_i}} = \mathrm{z}\right].$$

Numerical Calculation of $\Phi_i(z)$

➢ **Calculate $\Psi_i: \xi \mapsto \Phi_i\left(\xi T_i^{\frac{1}{2}}\right)$ on a set of N(0,1)-nodes $\{\xi_k: k = 1, \cdots, M\}$:**

Use $M = 2m$ and the symmetric Hermite nodes that appear in the Gauss-Hermite integration:

$$E[f(\xi)] = \sum_{k=1}^{M} p_k f(\xi_k).$$

The Hermite nodes associated with $M = 10$ are:

-3.775611	-2.516917	-1.50041	-0.682802	-0.141764	0.141764	0.682802	1.50041	2.516917	3.775611

➢ **Calculate Ψ_i by forward induction:**

$$\Psi_0(\xi) = 1,$$

$$\Psi_i(\xi_k) \approx E\left[\Psi_{i-1}(\zeta_{T_{i-1}}) \exp\left(\lambda \Delta T_i \frac{1}{\varphi\left(s_{i-1}\zeta_{T_{i-1}}\right)}\right) \middle| \zeta_{T_i} = \xi_k\right] \exp\left(\lambda \Delta T_i \frac{1}{\varphi(s_i \xi_k)}\right),$$

$$s_i = T_i^{\frac{1}{2}}, \qquad \lambda = \frac{3}{16}\gamma^2(1-\rho^2), \qquad \zeta_t = W_t t^{-\frac{1}{2}}.$$

The conditional expectation can be analytically computed since $\zeta_{T_{i-1}}$ and ζ_{T_i} are unit normal variables with correlation $\rho_i = (\frac{T_{i-1}}{T_i})^{\frac{1}{2}}$.

➢ **Represent the function $\boldsymbol{F_{i-1}(z) = \Psi_{i-1}(z) \exp\left(\lambda \Delta T_{i} \frac{1}{\varphi(s_{i-1}z)}\right)}$ by a cubic spline:**

$$\Psi_{i}(\xi_k) = E\left[F_{i-1}\left(\rho_i \xi_k + (1-\rho_i^2)^{\frac{1}{2}}\zeta\right)\right] \exp\left(\lambda \Delta T_{i} \frac{1}{\varphi(s_i \xi_k)}\right).$$

Express $F_{i-1}(z)$ using the associated basis spline functions:

$$F_{i-1}(z) = \sum_{j=1}^{M} F_{i-1}(\xi_j)\theta_j(z).$$

By integrating the above, simplify the former equation as follows:

$$\Psi_{i}(\xi_k) = \sum_{j=1}^{M} p_{kj}^{(i)} \times \Psi_{i-1}(\xi_j) \times \exp\left(\frac{\lambda \Delta T_{i}}{\varphi(s_{i-1}\xi_j)} + \frac{\lambda \Delta T_{i}}{\varphi(s_i \xi_k)}\right),$$

$$p_{kj}^{(i)} = p_{kj}(\rho_i) = E\left[\theta_j\left(\rho_i \xi_k + (1-\rho_i^2)^{\frac{1}{2}}\zeta\right)\right],$$

where $p_{kj}^{(i)}$ satisfies $\sum_j p_{kj}^{(i)} = 1$ but can be negative.

Observe that the pseudo-transition probabilities p_{kj} only depend on the mesh ξ_k and the grid T_i.

Consequently, the respective expectations only need to be computed once.

Analytical Calculation of the Pseudo-Transition Probabilities

The function θ_j is a cubic spline with value zero at every node except at $z = \xi_j$ where it takes value one:

$$\theta_j(z) = \sum_{i=1}^{M} 1\{L_i < z < U_i\} c_i(z),$$

$$L_{i\neq 0} = \xi_i, L_0 = -50, U_{i\neq M} = \xi_{i+1}, U_M = 50,$$

$$c_i(z) = A_{1i}(z - \xi_i)^3 + B_{1i}(z - \xi_i) - A_{0i}(z - \xi_{i+1})^3 - B_{0i}(z - \xi_{i+1}),$$

where $A_{0i}, B_{0i}, A_{1i}, B_{1i}$ are calculated using the standard cubic spline algorithm.

Finally, compute the pseudo-transition probabilities:

$$p_{kj}(\rho) = \sum_{i=0}^{M} \rho_\perp^3 \big(A_{1i} I_3(l_i, u_i, v_i) - A_{0i} I_3(l_i, u_i, v_{i+1})\big) + \rho_\perp (B_{1i}\, I_1(l_i, u_i, v_i) - B_{0i} I_1(l_i, u_i, v_{i+1})),$$

$$l_i = \frac{L_i - \rho\xi_k}{(1-\rho^2)^{\frac{1}{2}}}, \qquad u_i = \frac{U_i - \rho\xi_k}{(1-\rho^2)^{\frac{1}{2}}}, \qquad v_i = \frac{\rho\xi_k - \xi_i}{(1-\rho^2)^{\frac{1}{2}}}, \qquad \rho_\perp = (1-\rho^2)^{\frac{1}{2}},$$

where $I_n(a, b, c) = E[1\{a < \xi < b\}(\xi + c)^n]$.

By integration:

$$I_1(a, b, c) = c\big(N(b) - N(a)\big) + n(a) - n(b),$$

$$I_3(a, b, c) = (c^3 + 3c)\big(N(b) - N(a)\big) + (3c(c + a) + a^2 + 2)n(a) - (3c(c + b) + b^2 + 2)n(b).$$

Numerical Results

With standard market conditions, just a few steps *N* and states *M* are needed to obtain accurate call prices.

With $N = M = 10$, we obtain the following results where the implied volatility is displayed as a function of the lognormal standard deviation defined as $\frac{\ln K/S_0}{\sigma\sqrt{T}}$: