

Interest rate volatility

II. SABR and its flavors

Andrew Lesniewski

Baruch College and Posnania Inc

First Baruch Volatility Workshop
New York

June 16 - 18, 2015

Beyond local volatility models

- In general, local volatility models do not fit well the interest rate options prices.
- Among the issues is the “wing effect” exhibited by the implied volatilities of some maturities (especially short dated) and tenors which is not captured by these models: the implied volatilities tend to rise for high strikes forming the familiar “smile” shape.
- A way to address these issues is *stochastic volatility*. In this approach, a suitable volatility parameter is assumed to follow a stochastic process.
- The dynamics of the SABR model is given by:

$$\begin{aligned}dF(t) &= \sigma(t) C(F(t)) dW(t), \\d\sigma(t) &= \alpha\sigma(t) dZ(t).\end{aligned}\tag{1}$$

- Here F is the forward rate which, depending on context, may denote a LIBOR forward, a forward swap rate, or a forward bond yield¹, and σ is the volatility parameter.

¹The SABR model specification is also used in markets other than interest rate market, and thus F may denote e.g. a crude oil forward.

- The process is driven by two Brownian motions, $W(t)$ and $Z(t)$, with

$$E[dW(t) dZ(t)] = \rho dt, \quad (2)$$

where the correlation ρ is assumed constant.

- The diffusion coefficient $C(F)$ is assumed to be of the CEV type:

$$C(F) = F^\beta. \quad (3)$$

- The process $\sigma(t)$ is the stochastic component of the volatility of $F(t)$, and α is the volatility of $\sigma(t)$ (*vol of vol*), which is also assumed to be constant.
- We supplement the dynamics with the initial condition

$$\begin{aligned} F(0) &= F_0, \\ \sigma(0) &= \sigma_0, \end{aligned} \quad (4)$$

where F_0 is the current value of the forward, and σ_0 is the current value of the volatility parameter.

- Note that the dynamics (1) requires a boundary condition at $F = 0$. One usually imposes the absorbing (Dirichlet) boundary condition.

- It is conceptually and practically important that the process $F(t)$ generated by the SABR dynamics is a martingale.
- This issue is settled by a theorem proved by Jourdain [2]:
 - $F(t)$ is a martingale if $\beta < 1$ (it can be negative). If $\beta = 1$, $F(t)$ is a martingale only for $\rho < 0$.
 - Otherwise, $F(t)$ is not a martingale.

Monte Carlo simulation of SABR

- Despite its formal simplicity, the probability distribution associated with the SABR model is fairly complicated and can be accessed exactly only through Monte Carlo simulations.
- In order to assure that the stochastic volatility σ is positive, we rewrite the dynamics in terms of $X = \log \sigma$:

$$\begin{aligned}dF(t) &= \exp(X(t)) F(t)^\beta dW(t), \\dX(t) &= -\frac{1}{2} \alpha^2 dt + \alpha dZ(t).\end{aligned}\tag{5}$$

- This leads to the following *log-Euler scheme* based on the following discretization of the SABR dynamics

$$\begin{aligned}F_{k+1} &= (F_k + \sigma_k F_k^\beta \Delta W_k)^+, \\ \sigma_{k+1} &= \sigma_k \exp(\alpha \Delta Z_k - \delta \alpha^2 / 2),\end{aligned}\tag{6}$$

where δ is the time step, and $\Delta W_k, \Delta Z_k \sim N(0, \delta)$ are normal variates with variance δ and correlation coefficient ρ .

Monte Carlo simulation of SABR

- Note that the presence of $(\cdot)^+ = \max(\cdot, 0)$ imposes absorbing boundary condition at zero forward.
- The dynamics (5) is incompatible with the Milstein scheme: second order (in Brownian motion increments) discretization contains “Levy area” terms of the form $\int_t^{t+\delta} dW(s) dZ(s)$, which are hard to simulate.
- However, the following *quasi Milstein scheme*:

$$\begin{aligned} F_{k+1} &= (F_k + \sigma_k F_k^\beta \Delta W_k + \frac{\beta}{2} \sigma_k^2 F_k^{2\beta-1} (\Delta W_k^2 - \delta))^+, \\ \sigma_{k+1} &= \sigma_k \exp(\alpha \Delta Z_k - \delta \alpha^2 / 2), \end{aligned} \quad (7)$$

converges faster than the Euler scheme discussed above.

- The price of a European payer swaption is thus given by

$$P^{\text{pay}}(T, K, F_0) = A_0 \frac{1}{N} \sum_{1 \leq j \leq N} (F_n^{(j)} - K)^+, \quad (8)$$

where A_0 denotes the annuity function, and N is the number of Monte Carlo paths.

- Because of their relatively slow performance, Monte Carlo solutions are not always practical.
- A large portfolio of options at a broker / dealer, asset manager, or central counterparty requires frequent revaluations in order to:
 - recalibrate the model, as the market evolves,
 - update the portfolio risk metrics

These calculations require multiple applications of the pricing model.

- As a result, it is desirable to have a closed form solution of the model or at least a good analytic approximate solution.
- This requires a more detailed analysis of the model by means of asymptotic methods [5], [6].
- The use of asymptotic techniques in finance was pioneered in [4].

- The starting point of such an analysis is the terminal value problem for the backward Kolmogorov equation associated with the SABR process (1).
- We consider an “Arrow - Debreu security” with a singular payoff at expiration given by the Dirac function $\delta(F(T) - F)\delta(\sigma(T) - \Sigma)$.
- Its time t price $G = G(t, x, y; T, F, \Sigma)$ is called Green's function (where x corresponds to the forward and y corresponds to the volatility parameter), and is the solution to the following terminal value problem:

$$\frac{\partial}{\partial t} G + \frac{1}{2} y^2 \left(x^{2\beta} \frac{\partial^2}{\partial x^2} + 2\alpha\rho x^\beta \frac{\partial^2}{\partial x \partial y} + \alpha^2 \frac{\partial^2}{\partial y^2} \right) G = 0, \quad (9)$$
$$G(T, x, y; T, F, \Sigma) = \delta(x - F)\delta(y - \Sigma).$$

- Here, F and Σ are the terminal values of the forward and volatility parameter at option expiration T .

- The time $t < T$ price of a (call) option is then given by

$$P^{\text{call}}(T, K, F_0, \sigma) = \mathcal{N}(0) \int_0^\infty \int_0^\infty (F - K)^+ G(T - t, K, \sigma; F, \Sigma) d\Sigma dF. \quad (10)$$

- Note, in particular, that the terminal values of Σ are “integrated out”.
- From a numerical perspective, this expression is rather cumbersome: it requires solving the three dimensional PDE (9) and then calculating the double integral (10).
- Except for the special case of $\beta = 0$, no explicit solution to this model is known, and even in this case the explicit solution is too complex to be of practical use.
- We now outline how a practical, analytic solution can be constructed by means of asymptotic analysis.

Normal SABR model

- We first consider the normal SABR model [6], which is given by the following choice of parameters: $\beta = 0$ and $\rho = 0$.
- Equation (9) takes then the following form:

$$\begin{aligned} \frac{\partial}{\partial t} G + \frac{1}{2} y^2 \left(\frac{\partial^2}{\partial x^2} + \alpha^2 \frac{\partial^2}{\partial y^2} \right) G &= 0, \\ G(T, x, y; T, F, \Sigma) &= \delta(x - F) \delta(y - \Sigma). \end{aligned} \tag{11}$$

- For convenience, we change variables $\tau = T - t$ and rewrite the above terminal value problem as the initial value problem:

$$\begin{aligned} \frac{\partial}{\partial \tau} G &= \frac{1}{2} y^2 \left(\frac{\partial^2}{\partial x^2} + \alpha^2 \frac{\partial^2}{\partial y^2} \right) G, \\ G(0, x, y; 0, F, \Sigma) &= \delta(x - F) \delta(y - \Sigma). \end{aligned} \tag{12}$$

Normal SABR model

- This problem has an explicit solution given by *McKean's formula*:

$$G(\tau, x, y; F, \Sigma) = \frac{e^{-\alpha^2 \tau / 8} \sqrt{2}}{(2\pi \tau \alpha^2)^{3/2}} \int_d^\infty \frac{ue^{-u^2/2\alpha^2 \tau \alpha^2}}{\sqrt{\cosh u - \cosh d}} du. \quad (13)$$

- Here, $d = d(x, y, K, \Sigma)$ is the “geodesic distance” function given by

$$\cosh d(x, y, F, \Sigma) = 1 + \frac{\alpha^2(x - F)^2 + (y - \Sigma)^2}{2y\Sigma}. \quad (14)$$

- In order to proceed, *we assume* that the parameter $\varepsilon = \alpha^2 T$ is small. This will be the basis of the approximations that we make in the following. As it happens, this parameter is typically small for all swaptions and the approximate solution is quite accurate. A typical range of values of α is $0.2 \lesssim \alpha \lesssim 2$.²
- Also significantly, this solution is very easy to implement in computer code, and it lends itself well to risk management of large portfolios of options in real time.

²On a few days at the height of the recent financial crisis the value of α corresponding to 1 month into 1 year swaptions was as high as 4.7.

Normal SABR model

- We substitute $u = \sqrt{4\alpha^2\tau w + d^2}$ in the integral (13) in order to shift the lower limit of integration to 0. We then expand the integrand in a Taylor series in $\alpha^2\tau$.
- As a consequence, we obtain the following approximation. For $\tau \rightarrow 0$,

$$G(\tau, x, y; F, \Sigma) = \frac{1}{2\pi\alpha^2\tau} \sqrt{\frac{d}{\sinh d}} \exp\left(-\frac{d^2}{2\pi\tau\alpha^2}\right) \left(1 + O(\alpha^2\tau)\right).$$

- Next, we have to carry out the integration over the terminal values of Σ .
- To this end, we will use the steepest descent method: Assume that $\phi(x)$ is positive and has a unique minimum x_0 in $(0, \infty)$ with $\phi''(x_0) > 0$. Then, as $\varepsilon \rightarrow 0$,

$$\int_0^\infty f(x) e^{-\phi(x)/\varepsilon} du = \sqrt{\frac{2\pi\varepsilon}{\phi''(x_0)}} e^{-\phi(x_0)/\varepsilon} f(x_0) + O(\varepsilon). \quad (15)$$

- Specifically, the marginal density for the forward x is thus given by:

$$\begin{aligned} g(\tau, x, y; F) &= \int_0^\infty G(\tau, x, y; F, \Sigma) d\Sigma \\ &= \frac{1}{2\pi\alpha^2\tau} \int_0^\infty \sqrt{\frac{d}{\sinh d}} \exp\left(-\frac{d^2}{2\pi\tau\alpha^2}\right) d\Sigma + O(\alpha^2\tau). \end{aligned}$$

- We evaluate the integral above by means of the steepest descent method (15).
- The exponent $\phi(\Sigma) = \frac{1}{2} d(x, y, F, \Sigma)^2$ has a unique minimum at Σ_0 given by

$$\Sigma_0 = y\sqrt{\zeta^2 + 1},$$

where $\zeta = \alpha(x - F)/y$. Σ_0 is the “most likely value” of Σ , and thus it is the leading contribution to the observed implied volatility.

Normal SABR model

- Let $D(\zeta)$ denote the value of $d(x, y, K, \Sigma)$ with $\Sigma = \Sigma_0$. Explicitly,

$$D(\zeta) = \log(I(\zeta) + \zeta). \quad (16)$$

where we use the notation:

$$I(\zeta) = \sqrt{\zeta^2 + 1}. \quad (17)$$

- Then, we find that the terminal probability distribution is given, to within $O(\varepsilon)$, by

$$g(\tau, x, y; F) = \frac{1}{\sqrt{2\pi\tau}} \frac{1}{yI(\zeta)^{3/2}} \exp\left(-\frac{D^2}{2\tau\alpha^2}\right) \left(1 + O(\varepsilon)\right). \quad (18)$$

- Comparing this expression with the probability density of the normal distribution we see that

$$\sigma_n(T, K, F_0, \sigma_0, \alpha, \beta, \rho) = \alpha \frac{F_0 - K}{D(\alpha(F_0 - K)/\sigma_0)} \left(1 + O(\varepsilon)\right). \quad (19)$$

- This is the asymptotic expression for the implied normal volatility in the case of the normal SABR model.

Terminal probability distribution in the full SABR model

- The good news is that the full SABR model can *essentially* (up to some minor headaches) be mapped onto normal SABR
- This is accomplished by means of the following transformation of variables:

$$\begin{aligned}x' &= \frac{1}{\sqrt{1-\rho^2}} \left(\int_0^x \frac{dz}{C(z)} - \rho \frac{y}{\alpha} \right), \\ y' &= y.\end{aligned}\tag{20}$$

- This leads to the following expression, generalizing (15), for the terminal probability distribution:

$$\begin{aligned}g(\tau, x, y; F) &= \frac{1}{\sqrt{2\pi\tau}} \frac{1}{yC(F)I(\zeta)^{3/2}} \exp\left(-\frac{D(\zeta)^2}{2\tau\alpha^2}\right) \\ &\times \left(1 + \frac{yC'(x)D(\zeta)}{2\alpha\sqrt{1-\rho^2}I(\zeta)} + O(\varepsilon)\right).\end{aligned}\tag{21}$$

Terminal probability distribution in the full SABR model

- Here,

$$\begin{aligned}\zeta &= \frac{\alpha}{y} \int_F^x \frac{dz}{C(z)} \\ &= \frac{\alpha}{y(1-\beta)} (x^{1-\beta} - F^{1-\beta}).\end{aligned}\tag{22}$$

- The distance function $D(\zeta)$ is given by:

$$D(\zeta) = \log \left(\frac{I(\zeta) + \zeta - \rho}{1 - \rho} \right),\tag{23}$$

where

$$I(\zeta) = \sqrt{1 - 2\rho\zeta + \zeta^2}.\tag{24}$$

- These expressions reduce to the corresponding expressions that we derived for the normal SABR model when $\beta = 0$, $\rho = 0$.
- The most detailed analysis of the terminal probability distribution (up to the second order in ε) is carried out in [8].

Implied volatility in the full SABR model

- A careful analysis shows that the implied normal volatility in full SABR model is approximately given by:

$$\sigma_n(T, K, F_0, \sigma_0, \alpha, \beta, \rho) = \alpha \frac{F_0 - K}{D(\zeta)} \left\{ 1 + \left[\frac{2\gamma_2 - \gamma_1^2}{24} \right. \right. \\ \left. \left. \times \left(\frac{\sigma_0 C(F_{\text{mid}})}{\alpha} \right)^2 + \frac{\rho\gamma_1}{4} \frac{\sigma_0 C(F_{\text{mid}})}{\alpha} + \frac{2 - 3\rho^2}{24} \right] \varepsilon + \dots \right\}. \quad (25)$$

- Here, F_{mid} denotes a conveniently chosen midpoint between F_0 and K (such as $(F_0 + K)/2$), and

$$\gamma_1 = \frac{C'(F_{\text{mid}})}{C(F_{\text{mid}})},$$
$$\gamma_2 = \frac{C''(F_{\text{mid}})}{C(F_{\text{mid}})}.$$

Implied volatility in the full SABR model

- A similar asymptotic formula exists for the implied lognormal volatility σ_{\ln} . Namely,

$$\begin{aligned} \sigma_{\ln}(T, K, F_0, \sigma_0, \alpha, \beta, \rho) = & \alpha \frac{\log(F_0/K)}{D(\zeta)} \left\{ 1 + \left[\frac{2\gamma_2 - \gamma_1^2 + 1/F_{\text{mid}}^2}{24} \right. \right. \\ & \times \left. \left(\frac{\sigma_0 C(F_{\text{mid}})}{\alpha} \right)^2 + \frac{\rho\gamma_1}{4} \frac{\sigma_0 C(F_{\text{mid}})}{\alpha} + \frac{2 - 3\rho^2}{24} \right] \varepsilon + \dots \Big\}. \end{aligned} \quad (26)$$

Calibration of SABR

- For each option maturity we need to fix four model parameters: $\sigma_0, \alpha, \beta, \rho$. We choose them so that the model matches closely the market implied vols for several different strikes.
- It turns out that there is a bit of redundancy between the parameters β and ρ . As a result, one usually calibrates the model by fixing β .
- A popular choice is $\beta = 0.5$. This works quite well under “normal” conditions.
- In times of distress, such as during the crisis 2007 - 2009, the choice of $\beta = 0.5$ occasionally led to extreme calibrations of the correlation parameters ($\rho = \pm 1$). As a result, some practitioners choose high β 's, $\beta \approx 1$ for short expiry options and let it decay as option expiries move out.
- Calibration results show a persistent term structure of the model parameters as functions of option expiration and the underlying tenor. On a given market snapshot, the highest α is located in the upper left corner of the volatility matrix (short expirations and short tenors), and the lowest one is located in the lower right corner (long expirations and long tenors).

Calibration of SABR

- Figure 1 shows a market snapshot of SABR calibrated to the Eurodollar option volatilities.

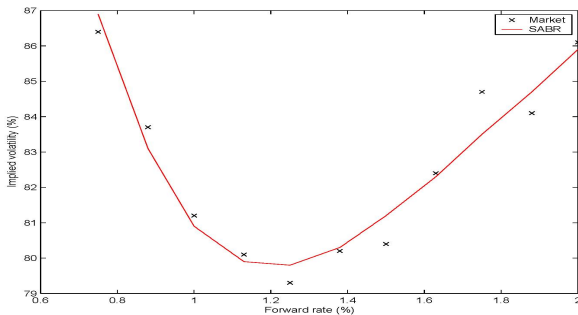


Figure: 1. SABR calibrated to ED options.

Calibration of SABR

- Graph 2 shows a snapshot of the expiry dependence of α .

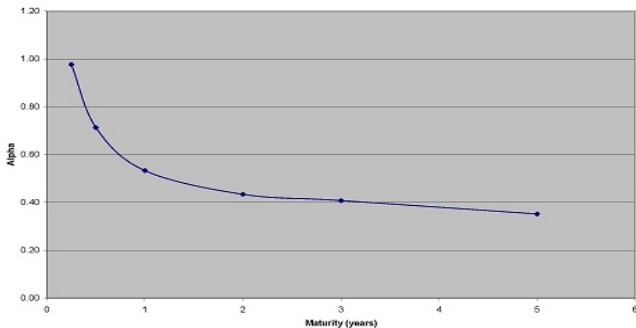


Figure: 2. Dependence of α on option expiration ($\beta = 0.5$)

Calibration of SABR

- Graph 3 shows a snapshot of the expiry dependence of ρ .

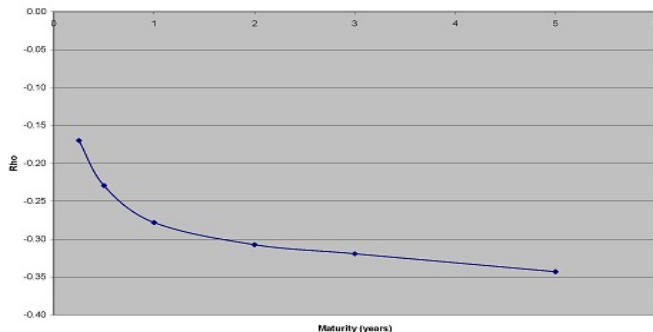


Figure: 3. Dependence of ρ on option expiration ($\beta = 0.5$)

Calibration of SABR

- Graph 4 shows the time series of the calibrated parameter α for the 5Y into 10Y swaption.

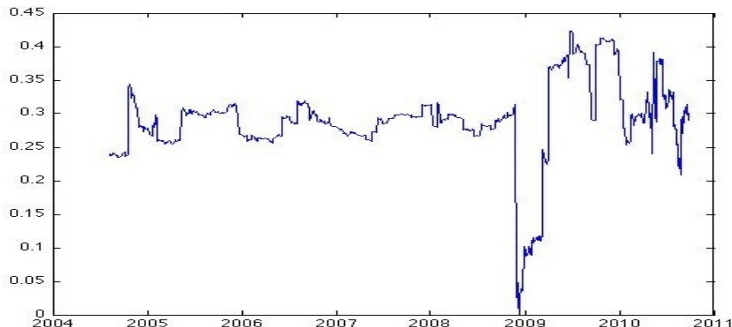


Figure: 4. Historical values of the calibrated parameter α ($\beta = 0.5$)

Calibration of SABR

- Graph 5 shows the time series of the calibrated parameter ρ for the 5Y into 10Y swaption. Notice the spike around the Lehman crisis.

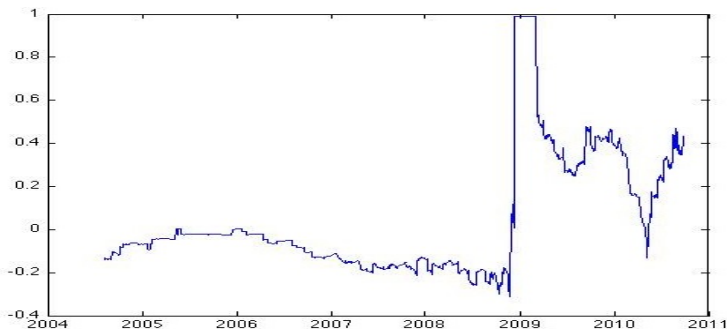


Figure: 5. Historical values of the calibrated parameter ρ ($\beta = 0.5$)

- There are two basic ways in which the asymptotic solution can be used for pricing options:
 - (i) based on the asymptotic terminal probability distribution, or
 - (ii) based on the asymptotic implied volatility formula.
- The former approach requires a numerical calculation of an integral of the form:

$$P^{\text{call}} = \mathcal{N}(0) \int_0^\infty (F - K)^+ g_T(F, F_0) dF, \quad (27)$$

where we use the notation $g_T(F, F_0) = g(T, F_0, \sigma_0; F)$.

- A potential issue with this approach is that the mean of the distribution $g_T(F, F_0)$ is not exactly F_0 and needs to be exogenously adjusted.

- The latter approach is more popular. We force the valuation formula to be of the form

$$\begin{aligned}P^{\text{call}}(T, K, F_0, \sigma_n) &= \mathcal{N}(0) B_n^{\text{call}}(T, K, F_0, \sigma_n), \\P^{\text{call}}(T, K, F_0, \sigma_n) &= \mathcal{N}(0) B_n^{\text{put}}(T, K, F_0, \sigma_n),\end{aligned}\tag{28}$$

given by the normal model, with the implied volatility $\sigma_n = \sigma_n(T, K, F_0, \sigma_0, \alpha, \beta, \rho)$ depending on the SABR model parameters.

- Under typical market conditions, these two approaches lead to identical results.

- Figure 6 shows the graphs of the asymptotic terminal PDF (black line) against the PDF implied by (28) (red line) in the case of a 3M option ($T = 0.25$, $\sigma_0 = 0.05$, $\alpha = 1.2$, $\beta = 0.5$, $\rho = -0.2$). The two graphs are very similar.

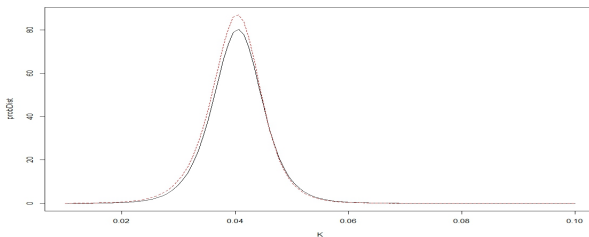


Figure: 6. Comparison of the asymptotic PDFs: short expiration

- Figure 6 shows the graphs of the asymptotic terminal PDF (black line) against the PDF implied by (28) (red line) in the case of a 5Y option ($T = 5$, $\sigma_0 = 0.05$, $\alpha = 0.4$, $\beta = 0.5$, $\rho = -0.2$). The differences between the two graphs are noticeable.

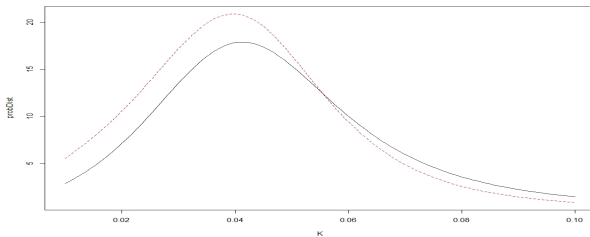


Figure: 7. Comparison of the asymptotic PDFs: long expiration

Adding mean reversion: λ -SABR

- An extension of SABR with mean reverting volatility parameter is given by the following system:

$$\begin{aligned}dF(t) &= \sigma(t) F(t)^{\beta} dW(t), \\d\sigma(t) &= \lambda(\mu - \sigma(t))dt + \alpha\sigma(t) dZ(t).\end{aligned}\tag{29}$$

Dealing with negative rates: shifted SABR

- Negative rates are abnormality but they are reality.
- In order to accommodate the SABR model to negative rates in the markets where they are observed (such as EUR), we shift the forward by a positive amount θ :

$$\begin{aligned}dF(t) &= \sigma(t) (F(t) + \theta)^\beta dW(t), \\ d\sigma(t) &= \alpha \sigma(t) dZ(t).\end{aligned}\tag{30}$$

- The shift θ cannot be directly calibrated to the market prices. A reasonable choice is $\theta = 4\%$.
- Explicit formulas for the implied volatilities and probability distributions are obtained by substituting $C(F) = (F + \theta)^\beta$ (in place of $C(F) = F^\beta$) in the corresponding equations in Presentation I.

Arbitrage freeness and SABR

- The explicit implied volatility given by formulas (25) or (26) make the SABR model easy to implement, calibrate, and use. These implied volatility formulas are usually treated *as if they were exact*, even though they are derived from an asymptotic expansion which requires that $\varepsilon = \alpha^2 T \ll 1$.
- The implicit assumption is that, instead of treating these formulas as an accurate approximation to the SABR model, they could be regarded as the exact solution to some other model which is well approximated by the SABR model. This is a valid viewpoint as long as the option prices obtained using the explicit formulas for σ_n (or σ_{ln}) are arbitrage free.
- There are two key requirements for arbitrage freeness of a volatility smile model:
 - (i) Put-call parity, which holds automatically since we are using the same implied volatility σ_n for both calls and puts.
 - (ii) The terminal probability density function implied by the call and put prices needs to be positive.

Arbitrage freeness and SABR

- To explore the second condition, recall that call and put prices can be written quite generally as

$$\begin{aligned}P^{\text{call}}(T, K) &= \mathcal{N}(0) \int_0^\infty (F - K)^+ g_T(F, F_0) dF, \\P^{\text{put}}(T, K) &= \mathcal{N}(0) \int_0^\infty (K - F)^+ g_T(F, F_0) dF,\end{aligned}\tag{31}$$

where $g_T(F, F_0)$ is the terminal PDF at the exercise date (possibly including the delta function from the Dirichlet boundary condition).

- As we saw in Presentation I,

$$\begin{aligned}\frac{\partial^2}{\partial K^2} P^{\text{call}}(T, K) &= \frac{\partial^2}{\partial K^2} P^{\text{put}}(T, K) \\&= g_T(K, F_0)\end{aligned}\tag{32}$$

- Arbitrage freeness is represented by the condition that

$$g_T(K, F_0) \geq 0,\tag{33}$$

for all K .

Arbitrage freeness and SABR

- In other words, there cannot be a “butterfly arbitrage”. As it turns out, it is not terribly uncommon for this requirement to be violated for very low strike and long expiry options. In the graph below, $T = 10$, $\sigma_0 = 0.05$, $\alpha = 0.1$, $\beta = 0.5$, $\rho = -0.2$.

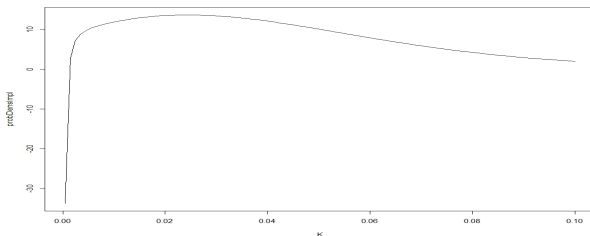


Figure: 8. Implied probability distribution of a 10Y option

Arbitrage freeness and SABR

- The problem does not appear to be the quality of the call and put prices obtained from the explicit implied volatility formulas, because these usually remain quite accurate.
- Rather, the problem seems to be that implied volatility curves are not a robust representation of option prices for low strikes. It is very easy to find a reasonable looking volatility curve $\sigma_n(K, \dots)$ which violates the arbitrage free constraint in (32) for a range of values of K .
- This issue is addressed by a number of authors, see [3], [1], and [7].

References

-  Andreasen, J., and Huge, B. N.: Expanded forward volatility, *Risk Magazine*, **January**, 101 - 107 (2013).
-  Jourdain, B.: Loss of martingality in asset price models with lognormal stochastic volatility, preprint (2014).
-  Doust, P.: No-arbitrage SABR, *J. Comp. Finance*, **15**, 3 - 31 (2012).
-  Hagan, P., and Woodward, D.: Equivalent Black volatilities, *Appl. Math. Finance*, **6**, 147 - 157 (1999).
-  Hagan, P., Kumar, D., Lesniewski, A., and Woodward, D.: Managing smile risk, *Wilmott Magazine*, **September**, 84 - 108 (2002).
-  Hagan, P., Lesniewski, A., and Woodward, D.: Probability distribution in the SABR model of stochastic volatility, preprint (2005).
-  Hagan, P., Kumar, D., Lesniewski, A., and Woodward, D.: Arbitrage Free SABR, *Wilmott Magazine*, **January** (2014).
-  Paulot, L.: Precision flying on the wings of SABR, preprint (2014).