

# Stochastic-alpha-beta-rho (SABR) Model

## Applied Stochastic Processes (FIN 514)

Instructor: Jaehyuk Choi

Peking University HSBC Business School, Shenzhen, China

2019-20 Module 1 (Fall 2019)

# The project overview

## SABR Model

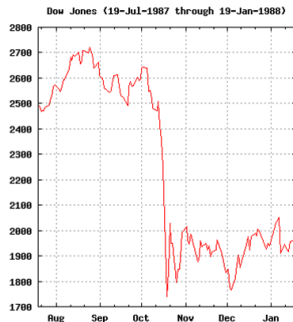
- One of the most popular **stochastic volatility (SV)** model.
- Heavily used for pricing and risk-managing options in interest rate and FX.
- Explains volatility skew/smile with minimal and intuitive parameters.

## Project Goal

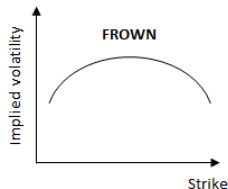
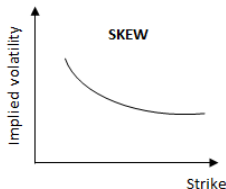
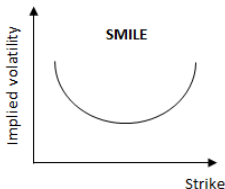
- Implement option pricing with Euler/Milstein scheme
- Implement conditional MC method (and check the variance reduction)
- Compare to the approximation formula by Hagan (code provided)
- Implement a smile calibration routine

# Background: volatility skew/smile

- Black Monday crash in 1987:  
DJIA  $-22.6\%$  in one day!
- Overall 'short gamma' due to the *portfolio insurance* (put on equity index)
- Market values (down-side) tail event higher than before.
- Market sees volatility skew/smile



(From Wikipedia)



# Why need model for smile? challenges in risk management

- Option trading desk (market-making/sell-side) usually accumulates option positions with different strikes.
- Under BSM model,
  - Vol  $\sigma$  fixed under spot change  $S_0 \rightarrow S_0 + \Delta$ .
  - Risk-management is easy: delta and vega clearly defined
  - One can hedge delta (with underlying stock) and vega (with ATM option)
  - However, **the OTM option prices/risks are not correct!**
- BSM model with different  $\sigma$  to each option  $K$ ?
  - How do we fix the volatilities?
  - Sticky strike rule  $\sigma = \sigma(K)$  vs sticky delta rule  $\sigma = \sigma(S_0 - K)$ .
  - Need to characterize the smile with a few minimal parameters.
- For better risk management, we need models which can capture the volatility smile.

# How to model smile? Local volatility (LV)

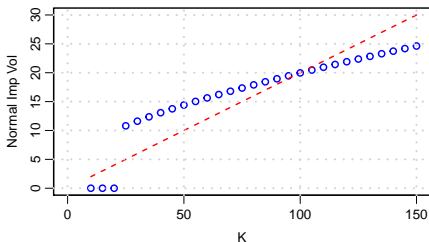
- Volatility depending on the 'current location' of  $S_t$ :

$$\text{BSM: } \frac{dS_t}{S_t} = \sigma f(S_t) dW_t \quad \text{Normal: } dS_t = \sigma_N f_N(S_t) dW_t$$

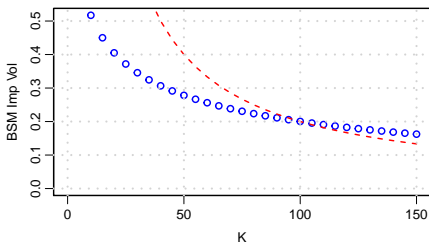
- BSM model:** a trivial case with  $f(x) = 1$ . However, it is a local vol model under normal volatility ( $f_N(x) = x$ ).
- Normal model:** a trivial case with  $f_N(x) = 1$ . However, it is a local vol model under BSM volatility ( $f(x) = 1/x$ ).
- What is the implied normal volatility of the Black-Scholes price on varying  $K$ ? What is the relation between the implied volatility and the local vol?
- The implied volatility is the volatility average of the in-the-money paths
- Exercise 1:** Chart the normal implied vol of the prices under BSM model for typical parameter sets. Measure the slope,  $\partial\sigma(K)/\partial K$ , at the money.

Case:  $S_0 = 100, \sigma = 20\% (\sigma_N = 20), r = q = 0$ :

- Implied normal vol for constant BSM vol ( $\sigma = 20\%$ ):



- Implied normal vol for constant normal vol ( $\sigma_N = 20$ ):



# Displaced GBM (shifted BSM) model

- A simple local vol model with analytic solution (i.e., Black-Scholes formula)
- *Displaced* (or *shifted*) asset price  $S_t + L$  follows a GBM:

$$dS_t = \sigma_L(S_t + L) dW_t$$

- Calibration of  $\sigma_L$  (ATM option price on target):

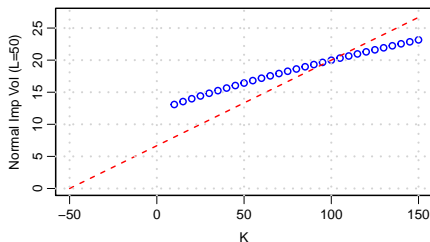
$$\sigma_N \approx \sigma_L(S_0 + L) \approx \sigma S_0$$

But, needs an exact calibration of  $\sigma_L$  for a given  $\sigma_{BS}$ .

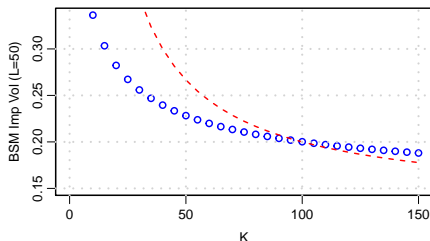
- Can reuse BS formula with  $S_0 + L \rightarrow S_0$  and  $K + L \rightarrow K$ .
- Somewhere between normal ( $L \rightarrow \infty$ ) and log-normal model ( $L = 0$ ).
- **Exercise 2:** Chart the BSM implied vol of the prices under displaced GBM model. Using the implemented implied vol function, exactly calibrate  $\sigma_L$  to the ATM price.

Case:  $S_0 = 100, L = 50, \sigma = 20\%, r = q = 0$ :

- $\sigma_L = \sigma S_0 / (S_0 + L) = 13.33\%$
- Implied normal vol: (red line:  $\sigma_L(K + L)$ )



- Implied BSM vol: (red line:  $\sigma_L(K + L)/K$ )





# How to model smile? Stochastic volatility (SV)

- Volatility changing over time:

$$\text{BSM: } \frac{dS_t}{S_t} = \sigma_t dW_t \quad \text{Normal: } dS_t = \sigma_t dW_t$$

- Many models proposed (mostly for BSM). For  $dW_t dZ_t = \rho dt$ ,
  - Hull-White (SABR):

$$\frac{d\sigma_t}{\sigma_t} = \alpha dZ_t$$

- Heston:  $V_t = \sigma_t^2$  follows Cox-Ingersoll-Ross (CIR) process,

$$dV_t = \kappa(V_\infty - V_t)dt + \alpha\sqrt{V_t}dZ_t$$

- SV model correctly captures the smile,  $\alpha$  for curvature and  $\rho$  for skewness.

Stochastic- $\alpha, \beta, \rho$  model SDE:

$$dS_t = \sigma_t S_t^\beta dW_t$$

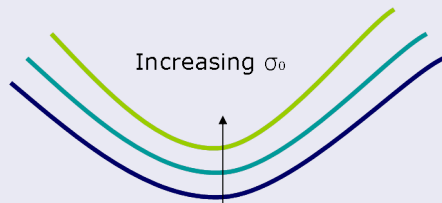
$$d\sigma_t = \alpha \sigma_t dZ_t$$

$$dW_t dZ_t = \rho dt$$

- Parameters:  $\sigma_0, \alpha, \beta, \rho$ .
- $\sigma_0$ : overall volatility, calibrated to ATM implied vol
- $\beta$ : elasticity or 'backbone'. (Normal:  $\beta = 0$ , BSM:  $\beta = 1$ )
- $\alpha$ : volatility of volatility,  $\sigma$  following a GBM
- $\rho$ : correlation between asset price and volatility

# The impact of parameters

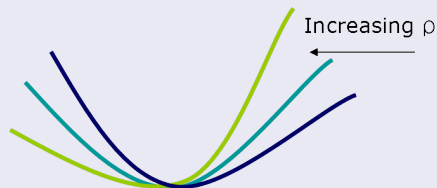
## Initial vol $\sigma_0$



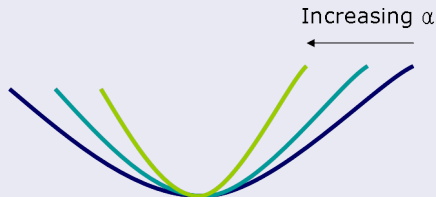
## Backbone $\beta$

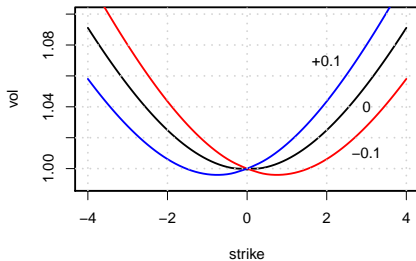
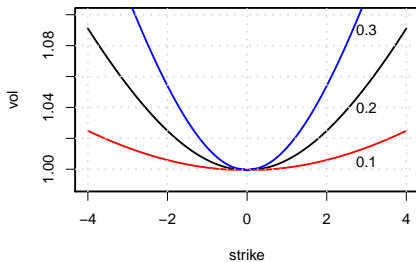
- Fixed or infrequently changed
- BSM model:  $\beta = 1$  (Equity, FX)
- Normal model:  $\beta = 0$  (Interest Rate)

## Correlation $\rho$



## 'Vol of vol' $\alpha$





# Equivalent BSM-volatility formula (Hagan et al, 2002)

The first few terms of Taylor's expansion near  $\alpha\sqrt{T} \approx 0$ .

$$\begin{aligned}\sigma_B(K, f) &= \frac{\alpha}{(fK)^{(1-\beta)/2} \left\{ 1 + \frac{(1-\beta)^2}{24} \log^2 f/K + \frac{(1-\beta)^4}{1920} \log^4 f/K + \dots \right\}} \cdot \left( \frac{z}{x(z)} \right) \\ &\cdot \left\{ 1 + \left[ \frac{(1-\beta)^2}{24} \frac{\alpha^2}{(fK)^{1-\beta}} + \frac{1}{4} \frac{\rho\beta v\alpha}{(fK)^{(1-\beta)/2}} + \frac{2-3\rho^2}{24} v^2 \right] t_{ex} + \dots \right\}.\end{aligned}\tag{2.17a}$$

Here

$$z = \frac{v}{\alpha} (fK)^{(1-\beta)/2} \log f/K,\tag{2.17b}$$

and  $x(z)$  is defined by

$$x(z) = \log \left\{ \frac{\sqrt{1 - 2\rho z + z^2} + z - \rho}{1 - \rho} \right\}.\tag{2.17c}$$

# Success of the SABR model

- Volatility smile information encoded into three parameters  $\sigma_0, \alpha, \rho$ .
- These three parameters are parsimonious (minimal) and intuitive.
- Equivalent BSM volatility is available although not accurate for wide parameter range.
- Vega (volatility) risk managed by the three parameters rather than each individual vol.
- Three implied vols (or option prices) on the smile can calibrate the parameters. → An effective interpolation method for implied volatility (or option price)

# Limitation of Hagan's formula

- Arbitrage is equivalent to some event happening with negative probability. The price of a derivative paying \$1 on that event is negative (should be free at most)!
- Digital call option price (probability) from call spread:

$$\begin{aligned} P(S_T > K) &= D(K, \sigma(K)) \\ &= \frac{C_{BS}(K, \sigma(K)) - C_{BS}(K + \Delta K, \sigma(K + \Delta K))}{\Delta K} = - \frac{\partial C_{BS}(K, \sigma(K))}{\partial K} \end{aligned}$$

- For positive PDF,  $D(K, \sigma(K))$  should be monotonically decreasing on  $K$ . When  $\alpha\sqrt{T} \gg 1$ , however, Hagan's formula often implies  $D(K, \sigma(K))$  increasing on  $K$ :

$$D(K, \sigma(K)) < D(K + \Delta K, \sigma(K + \Delta K)).$$

The volatility effect  $\sigma(K + \Delta K)$  overcomes (should NOT!) the moneyness effect  $K + \Delta K$ .

# Euler method (MC with time-discretization)

- Unlike normal or BSM model (as in spread/basket option project), we can not jump the simulation directly from  $t = 0$  to  $T$ .
- Divide the interval  $[0, T]$  into  $N$  small steps,  $t_k = (k/N)T$  and  $\Delta t_k = T/N$  and simulate each time step with

$$S_t : \begin{cases} \beta = 0 : S_{t_{k+1}} = S_{t_k} + \sigma_{t_k} W_1 \sqrt{\Delta t_k} \\ \beta = 1 : \log S_{t_{k+1}} = \log S_{t_k} + \sigma_{t_k} \sqrt{\Delta t_k} W_1 - \frac{1}{2} \sigma_{t_k}^2 \Delta t_k, \end{cases}$$
$$\sigma_t : \sigma_{t_{k+1}} = \sigma_{t_k} \exp \left( \alpha \sqrt{\Delta t_k} Z_1 - \frac{1}{2} \alpha^2 \Delta t_k \right),$$

where  $W_1, Z_1 \sim N(0, 1)$  with correlation  $\rho$ .

- Typically,  $\Delta t_k \approx 0.25$ . For  $T = 30$ ,  $N = 120$ , quite time-consuming.
- Any good control variate?

$$C(K) = \frac{1}{N} \sum_{i=1}^N (S_T^{(i)} - K)^+$$



# Euler method vs Milstein method

For a stochastic process,

$$dX_t = \mu(X_t)dt + \sigma(X_t)dW_t,$$

the Euler scheme is given as:

$$X_{t+\Delta t} = X_t + \mu(X_t)\Delta t + \sigma(X_t)\Delta W_t$$

In Milstein scheme, an higher-order correction is added:

$$X_{t+\Delta t} = X_t + \mu(X_t)\Delta t + \sigma(X_t)\Delta W_t + \frac{\sigma(X_t)\sigma'(X_t)}{2}((\Delta W_t)^2 - \Delta t),$$

where the extra term is derived from the GBM approximation:

$$\sigma(X_t) = A X_t \quad \text{with} \quad A = \sigma'(X_t) \quad \text{and} \quad B = \mu(X_t)/X_t$$

$$d \log X_t = (B - \frac{1}{2}A^2)dt + A dW_t,$$

$$X_{t+\Delta t} = X_t \cdot \exp \left( (B - \frac{1}{2}A^2)\Delta t + A \Delta W_t \right)$$

# Stochastic integral of $\sigma_t$

From Itô's lemma,

$$\frac{d\sigma_t}{\sigma_t} = \alpha dZ_t \quad \Rightarrow \quad d \log \sigma_t = -\frac{1}{2}\alpha^2 dt + \alpha dZ_t$$

we can solve the volatility process:

$$\sigma_T = \sigma_0 \exp \left( -\frac{1}{2}\alpha^2 T + \alpha Z_T \right).$$

We also know

$$\alpha \int_0^T \sigma_t dZ_t = \sigma_T - \sigma_0 = \sigma_0 \exp \left( -\frac{1}{2}\alpha^2 T + \alpha Z_T \right) - \sigma_0,$$

which will be useful for the integration of  $S_t$ .

# Stochastic integral of $S_t$ (normal: $\beta = 0$ )

Writing the SDE in a de-correlated form,

$$dS_t = \sigma_t \left( \rho dZ_t + \sqrt{1 - \rho^2} dX_t \right) \quad \text{with} \quad dX_t dZ_t = 0.$$

Integrating  $S_t$ , we get so far as

$$\begin{aligned} S_T - S_0 &= \rho \int_0^T \sigma_t dZ_t + \sqrt{1 - \rho^2} \int_0^T \sigma_t dX_t \\ &= \frac{\rho}{\alpha} (\sigma_T - \sigma_0) + \sqrt{1 - \rho^2} \int_0^T \sigma_t dX_t \end{aligned}$$

From Itô's Isometry, the integration in blue is equivalent to

$$\int_0^T \sigma_t dX_t = X_1 \sqrt{I_T} \quad \text{where} \quad X_1 \sim N(0, 1), \quad I_T := \int_0^T \sigma_t^2 dt.$$

Here, the random variable  $X_1$  is independent from  $I_T$  and  $\sigma_T$ .

# Conditional MC method (normal $\beta = 0$ )

Conditional on  $(\sigma_T, I_T)$ , the distribution of  $S_T$  is

$$S_T = S_0 + \frac{\rho}{\alpha}(\sigma_T - \sigma_0) + \sqrt{(1 - \rho^2)I_T} X_1$$

and the option price is from the normal model:

$$C_N \left( K, S_0 := S_0 + \frac{\rho}{\alpha}(\sigma_T - \sigma_0), \sigma_N := \sqrt{(1 - \rho^2)I_T/T} \right)$$

Then, the price is obtained as an expectation over MC simulation of  $I_T$ :

$$C_{\beta=0} = E(C_N(\sigma_T, I_T)), \quad \text{where} \quad I_T = \sum_k \sigma_{t_k}^2 \Delta t_k$$

In this way, no need for simulating  $S_t$ . Given  $(\sigma_T, I_T)$ , the option price is exact. Therefore, MC variance is much smaller than that of the MC simulating both  $\sigma_t$  and  $S_t$ .

# Conditional MC method (BSM $\beta = 1$ )

Conditional on  $(\sigma_T, I_T)$ , the distribution of  $S_T$  is

$$\log \left( \frac{S_T}{S_0} \right) = \frac{\rho}{\alpha} (\sigma_T - \sigma_0) - \frac{1}{2} I_T + \sqrt{1 - \rho^2} X_1 \sqrt{I_T}$$

and the option price is from the BSM formula:

$$C_{\text{BS}} \left( K, S_0 e^{\frac{\rho}{\alpha} (\sigma_T - \sigma_0) - \frac{1}{2} I_T}, \sqrt{(1 - \rho^2) I_T / T} \right)$$

Then, the price is obtained as an expectation over MC simulation of  $I_T$ :

$$C_{\beta=1} = E(C_{\text{BS}}(\sigma_T, I_T)), \quad \text{where} \quad I_T = \sum_k \sigma_{t_k}^2 \Delta t_k$$

In this way, no need for simulating  $S_t$ . Given  $(\sigma_T, I_T)$  the price is exact, therefore MC variance is much smaller than that of brute-force MC.

# The conditional distribution of $I_T$ on $\sigma_T$ (Kennedy et al)

The conditional mean of  $I_T$  on  $\sigma_T$  is known as

$$E(I_T|\sigma_T) = \frac{\sigma_0^2 \sqrt{T}}{2\alpha} \frac{N(d_\alpha + \alpha\sqrt{T}) - N(d_\alpha - \alpha\sqrt{T})}{n(d_\alpha + \alpha\sqrt{T})}$$

for  $d_\alpha = \log(\sigma_T/\sigma_0)/(\alpha\sqrt{T})$ .

The distribution of  $S_T$  is approximated

$$S_T = S_0 + \frac{\rho}{\alpha}(\sigma_T - \sigma_0) + \sqrt{1 - \rho^2} \eta(\sigma_T) W \sqrt{T}$$

for  $\eta(\sigma_T) = E(I_T|\sigma_T)/\sqrt{T}$ . For a given  $\sigma_T$ ,  $S_T$  follows a normal distribution, so we now the option

$$C_N(\sigma_T) = C_N \left( S_0 := S_0 + \frac{\rho}{\alpha}(\sigma_T - \sigma_0), \sigma_N := \sqrt{1 - \rho^2} \eta(\sigma_T) \right)$$

# Option price as an integration (Kennedy et al)

$$\begin{aligned}C_{\beta=0} &= E\left((S_T - K)^+\right) = E\left((S_T - K)^+ | \sigma_T\right) = E\left(C_N(\sigma_T)\right) \\&= \int_{-\infty}^{\infty} C_N\left(S_0 + \frac{\rho}{\alpha}(\sigma_T(z) - \sigma_0), \sqrt{1 - \rho^2} \eta(\sigma_T(z))\right) n(z) dz \\&\quad \text{where } \sigma_T(z) = \sigma_0 \exp\left(-\frac{1}{2}\alpha^2 T + \alpha\sqrt{T} z\right)\end{aligned}$$

Using Gauss-Hermite quadrature (GHQ), [Py Demo]

$$C_{\beta=0} = \sum_m C_N\left(S_0 + \frac{\rho}{\alpha}(\sigma_T(z_m) - \sigma_0), \sqrt{1 - \rho^2} \eta(z_m)\right) w_m$$

for some points  $\{z_m\}$  and weights  $\{w_m\}$ , and  $\eta(z_m) := \eta(\sigma_T(z_m))$ .

The results are similar:

$$\log(S_T/S_0) = \frac{\rho}{\alpha}(\sigma_T - \sigma_0) - \frac{1}{2}I_T + \sqrt{1 - \rho^2} W \sqrt{I_T}$$

$$C_{\beta=1} = \sum_m C_{BS} \left( S_0 e^{\frac{\rho}{\alpha}(\sigma_T(z_m) - \sigma_0) - \frac{\rho^2}{2}\eta(z_m)}, \sqrt{1 - \rho^2} \eta(z_m) \right) w_m$$

for some points  $\{z_m\}$  and weights  $\{w_m\}$ .

- Implement the method of Kennedy et al and compare it against the Monte Carlo result for both normal ( $\beta = 0$ ) and BSM backbone ( $\beta = 1$ ).



- When  $\beta$  is given (0 or 1), three parameters,  $\sigma_0$ ,  $\rho$  and  $\alpha$ , can be calibrated to three option prices (or implied volatilities), typically at  $K = S_0$  (ATM),  $S_0 - \Delta$  and  $S_0 + \Delta$ .

$$\text{SABR}(\sigma_0, \rho, \alpha) \rightarrow \sigma(S_0), \sigma(S_0 - \Delta), \sigma(S_0 + \Delta)$$

- Write a calibration routine in R to solve  $\sigma_0$ ,  $\rho$  and  $\alpha$ .

## Validation

- $C_{\beta=0}$  should converge to the normal model price when  $\alpha$  is very small.
- Test against the result from Korn & Tang (Wilmott)

## Homework

- First focus on the normal backbone  $\beta = 0$ .
- Make sure to use antithetic method (create  $Z$ , then add  $-Z$ ).
- Short (1 page) write-up briefly explaining the code.
- Reproduce the graphs in 'The impact of parameters'. Fix your normal implied vol at ATM.
- Your script should be self-complete and should run without error.