Interest Rate and Credit Models

5. Caps, Floors, and Swaptions

Andrew Lesniewski

Baruch College
New York

Spring 2019

Outline

Options on LIBOR based instruments

- Interest rates fluctuate as a consequence of macroeconomic conditions, central bank actions, and supply and demand.
- The existence of the *term structure of rates*, i.e. the fact that the level of a rate depends on the maturity of the underlying loan, makes the dynamics of rates is highly complex.
- While a good analogy to the price dynamics of an equity is a particle moving in a medium exerting random shocks to it, a natural way of thinking about the evolution of rates is that of a string moving in a random environment where the shocks can hit any location along its length.

Options on LIBOR based instruments

- Additional complications arise from the presence of various spreads between rates, as discussed in Lecture Notes #1, which reflect credit quality of the borrowing entity, liquidity of the instrument, or other market conditions.
- We will describe some of the mathematical methodologies to capture and quantify this dynamics that are widely used to:
  (i) assign a value on the optionality embedded in various financial instruments,
  (ii) help risk manage portfolios of fixed income securities,
  (iii) help identify mispricings and trading opportunities in the fixed income markets.

Options on LIBOR based instruments

- We have already taken the first step in this direction, namely learned how to construct the current snapshot of the rates market.
- This current snapshot serves as the initial value for the stochastic process describing the curve dynamics.
- The next step is to construct the *volatility cube*, which is used to model the uncertainties in the future evolution of the rates.
- The volatility cube is built out of implied volatilities of a number of liquidly trading options.
- Note that the presence of the third dimension in the volatility object[1] is another consequence of the presence of the term structure of interest rates.

[1] as compare to the volatility surface familiar from the world of equity derivatives

Caps and floors

- *Caps* and *floors* are baskets of European calls (called *caplets*) and puts (called *floorlets*) on LIBOR forward rates. They trade over the counter.
- Let us consider, for example, a 10 year spot starting cap struck at 2.50%. It consists of 39 caplets each of which expires on the 3 month anniversary of today's date.
- A caplet pays

$$\max(\text{current LIBOR fixing} - 2.50\%, 0) \times \text{act/360 day count fraction.}$$

  The payment is made at the end of the 3 month period covered by the LIBOR contract, and follows the modified business day convention.
- Notice that the very first period is excluded from the cap: this is because the current LIBOR fixing is already known and no optionality is left in that period.

Caps and floors

- In addition to spot starting caps and floors, *forward starting* instruments are traded.
- For example, a 1 year $\times$ 5 years (in the market lingo: "1 by 5") cap struck at 2.50% consists of 16 caplets struck at 2.50%, the first of which matures one year from today.
- The final maturity of the contract is 5 years, meaning that the last caplets matures 4 years and 9 months from today (with appropriate business dates adjustments).
- Unlike the case of spot starting caps, the first period is included into the structure, as the first LIBOR fixing is unknown. Note that the total maturity of the $m \times n$ cap is n years.
- The definitions of floors are similar with the understanding that a floorlet pays

$$\max(\text{strike} - \text{current LIBOR fixing}, 0) \times \text{act/360 day count fraction}$$

at the end of the corresponding period.

Eurodollar options

- *Eurodollar options* are standardized contracts traded at the Merc.
- These are short dated American style calls and puts on Eurodollar futures.
- At each time, options on the first eight (Fronts and Reds) quarterly Eurodollar futures contracts and on two front serial futures are listed.
- Their expirations coincide with the maturity dates of the underlying Eurodollar contracts.
- The exchange sets the strikes for the options spaced every 25 basis points (or 12.5 bp for the front contracts).
- The options are cash settled.

Eurodollar options

- Below is a snapshot of the ED options market:

Strike	Calls	Puts
98.875	0.5325	0.0525
99.000	0.4175	0.0625
99.125	0.3075	0.0775
99.250	0.2025	0.0975
99.375	0.1125	0.1325
99.500	0.0450	0.1900
99.625	0.0100	0.2800
99.750	0.0025	0.3975
99.875	0.0025	0.5200

Table: 1. ED options: March 2012 expirations. Price of the underlying 99.355

Eurodollar options

- In addition to the quarterly and serial contracts, a number of *midcurve* options are listed on the Merc.
- These are American style calls and puts with expirations between three months and one year on longer dated Eurodollar futures.
- Their expirations do not coincide with the maturity on the underlying futures contracts, which mature one, two, or four years later.
- The prices of all Eurodollar options are quoted in *ticks*.

Eurodollar options

- Below is a snapshot of the ED midcurve options market:

Strike	Calls	Puts
98.875	0.5275	0.0925
99.000	0.4200	0.1100
99.125	0.3150	0.1300
99.250	0.2175	0.1575
99.375	0.1275	0.1925
99.500	0.0650	0.2250
99.625	0.0250	0.3400
99.750	0.0075	0.4475
99.875	0.0025	0.5650

Table: 2. ED options: June 2012 expirations. Price of the underlying 99.31

Swaptions

- European *swaptions* are European calls and puts on interest rate swaps.
- In the market lingo calls and puts on swaps are called *receivers* and *payers*, respectively.
- A holder of a payer swaption has the right, upon exercise, to pay fixed coupon on a swap of contractually defined terms.
- Likewise, a holder of a receiver swaption has the right to receive fixed on a swap.
- Swaptions are traded over the counter.
- Swaptions with expirations up to 2 years are centrally cleared by CME.

Swaptions

- For example, a 2.50% 1Y $\to$ 5Y (“1 into 5”) receiver swaption gives the holder the right to receive 2.50% on a 5 year swap starting in 1 year.
- More precisely, the option holder has the right to exercise the option on the 1 year anniversary of today (with the usual business day convention adjustments) in which case they enter into a receiver swap starting two business days thereafter.
- Similarly, a 3.50% 5Y $\to$ 10Y (“5 into 10”) payer swaption gives the holder the right to pay 3.50% on a 10 year swap starting in 5 years.
- Note that the total maturity of the $m \to n$ swaption is $m + n$ years.

Swaptions

- Table 3 contains a snapshot of the at the money swaption market.
- The rows in the matrix represent the swaption expiration and the columns represent the tenor of the underlying swap.
- Each entry in the table represents the swaption premium expressed as a percentage of the notional on the underlying swap.

Swaptions

	1Y	2Y	3Y	4Y	5Y	7Y	10Y	15Y	20Y	30Y
1M	0.06%	0.11%	0.18%	0.27%	0.37%	0.67%	1.10%	1.70%	2.17%	2.94%
3M	0.10%	0.20%	0.31%	0.48%	0.68%	1.18%	1.91%	2.90%	3.69%	5.02%
6M	0.14%	0.30%	0.47%	0.74%	1.04%	1.73%	2.71%	4.06%	5.17%	6.97%
1Y	0.21%	0.45%	0.75%	1.16%	1.60%	2.51%	3.82%	5.56%	7.05%	9.45%
2Y	0.40%	0.85%	1.37%	1.94%	2.55%	3.66%	5.26%	7.38%	9.23%	12.20%
3Y	0.62%	1.26%	1.91%	2.58%	3.25%	4.50%	6.26%	8.61%	10.64%	13.77%
4Y	0.78%	1.54%	2.28%	3.02%	3.75%	5.11%	7.00%	9.52%	11.66%	15.11%
5Y	0.88%	1.74%	2.56%	3.35%	4.13%	5.58%	7.57%	10.21%	12.49%	16.15%
7Y	0.97%	1.90%	2.78%	3.63%	4.44%	5.97%	8.09%	10.81%	13.16%	16.86%
10Y	1.01%	1.96%	2.86%	3.71%	4.53%	6.08%	8.22%	10.86%	13.12%	16.71%

Table: 3. ATM swaption prices

Swaptions

- Since a swap can be viewed as a particular basket of underlying LIBOR forwards, a swaption is an option on a basket of forwards.
- This observation leads to the popular relative value trade of, say, a $2 \to 3$ swaption straddle versus a 2×5 cap / floor straddle.
- Such a trade my reflect the trader's view on the correlations between the LIBOR forwards or a misalignment of swaption and cap / floor volatilities.
- In addition to the standard swaptions discussed above, there is considerable liquidity in *midcurve swaptions*.
- The swap underlying a midcurve swaption starts on a date later than the spot date following the exercise date (say, 6 months later).

Black's model

- The market standard for quoting prices on caps / floors and swaptions is in terms of *Black's model*. This is a version of the Black-Scholes model adapted to handle forward underlying assets.
- We will now briefly discuss this model, and then describe some popular extensions of Black's model.
- We assume that a forward rate $F(t)$, such as a LIBOR forward or a forward swap rate, follows a driftless lognormal process reminiscent of the basic Black-Scholes model,
$$dF(t) = \sigma F(t)\, dW(t)\,. \tag{1}$$
Here $W(t)$ is a Wiener process, and σ is the *lognormal volatility*.
- It is understood here, that we have chosen an appropriate numeraire $\mathcal{N}$, depending on the instrument.
- We let Q denote the corresponding EMM.

Black's model

- The solution to this stochastic differential equation reads:

$$F(t) = F_0 \exp\left(\sigma W(t) - \frac{1}{2}\sigma^2 t\right). \tag{2}$$

- We consider a European call struck at K and expiring in T years.
- Assuming that the numeraire has been chosen so that $\mathcal{N}(T) = 1$, we can write its today's value as

$$\begin{aligned} P^{\text{call}}(T, K, F_0, \sigma) &= \mathcal{N}(0)\,\mathrm{E}^{\mathrm{Q}}\left[\max\left(F(T) - K, 0\right)\right] \\ &= \mathcal{N}(0)\,\frac{1}{\sqrt{2\pi T}}\int_{-\infty}^{\infty}\max\left(F_0 e^{\sigma W - \frac{1}{2}\sigma^2 T} - K, 0\right) e^{-\frac{W^2}{2T}}\,dW, \end{aligned} \tag{3}$$

where E^{Q} denotes expected value with respect to Q.

Black's model

- The last integral can easily be carried out, and we find that

$$\begin{aligned} P^{\text{call}}(T, K, F_0, \sigma) &= \mathcal{N}(0)\left[F_0 N(d_+) - K N(d_-)\right] \\ &\equiv \mathcal{N}(0)\, B_{\text{call}}(T, K, F_0, \sigma)\,. \end{aligned} \tag{4}$$

- Here,

$$N(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-y^2/2} dy \tag{5}$$

is the cumulative normal distribution, and

$$d_\pm = \frac{\log \frac{F_0}{K} \pm \frac{1}{2}\sigma^2 T}{\sigma\sqrt{T}}\,. \tag{6}$$

Black's model

- Similarly, the price of a European put is given by:

$$\begin{aligned} P^{\text{put}}(T, K, F_0, \sigma) &= \mathcal{N}(0)\left[-F_0 N(-d_+) + K N(-d_-)\right] \\ &\equiv \mathcal{N}(0)\, B_{\text{put}}(T, K, F_0, \sigma). \end{aligned} \tag{7}$$

Valuation of caps and floors

- A cap is a basket of options on LIBOR forward rates. Consider the OIS forward rate $F(S, T)$ spanning the accrual period $[S, T]$.
- Its time $t \leq S$ value $F(t, S, T)$ can be expressed in terms of discount factors:

$$\begin{aligned} F(t, S, T) &= \frac{1}{\delta} \left(\frac{P(t, t, S)}{P(t, t, T)} - 1 \right) \\ &= \frac{1}{\delta} \, \frac{P(t, t, S) - P(t, t, T)}{P(t, t, T)} \, . \end{aligned} \tag{8}$$

Valuation of caps and floors

- The interpretation of this identity is that $F(t, S, T)$ is a tradable asset if we use the zero coupon bond maturing in T years as numeraire.
- Indeed, the trade is as follows:
  - (i) Buy $1/\delta$ face value of the zero coupon bond for maturity S.
  - (ii) Sell $1/\delta$ face value of the zero coupon bond for maturity T.
- The value of this position in the units of $P(t, t, T)$ is $F(t, S, T)$. An OIS forward rate can thus be modeled as a martingale!
- We call the corresponding martingale measure the *T-forward measure* and denote it by Q_T.

Deterministic basis assumption

- Consider now a LIBOR forward $L(S, T)$ spanning the same accrual period.
- *Throughout these lectures we will make the assumption that the LIBOR / OIS spread is deterministic* (rather than stochastic).
- This assumption is, clearly, a gross oversimplification of reality but it has some merits.
- There are no liquidly trading options on this spread, and thus calibrating a model with a stochastic spread is problematic.
- From the conceptual point of view, the picture is more transparent with a deterministic spread. Namely, we know from Lecture Notes #1 that

$$L(t, S, T) = F(t, S, T) + B(S, T). \tag{9}$$

- This shows that the LIBOR forward is a martingale under the T-forward measure Q_T.

Valuation of caps and floors

- A caplet is a put option on the zero coupon bond. Let K be the strike on the caplet. Then, under the martingale measure Q_T,

$$
\begin{aligned}
&\delta \mathsf{E}^{\mathsf{Q}_T}\Big[\Big(L(S,S,T)-K\Big)^+\Big] \\
&\qquad = \mathsf{E}^{\mathsf{Q}_T}\Big[\Big(\frac{1-(1-\delta B(S,S,T))P(S,S,T)}{P(S,S,T)}-\delta K\Big)^+\Big] \\
&\qquad = \frac{1}{\widetilde{K}}\,\mathsf{E}^{\mathsf{Q}_T}\Big[\frac{(\widetilde{K}-P(S,S,T))^+}{P(S,S,T)}\Big],
\end{aligned}
$$

where

$$
\widetilde{K} = \frac{1}{1+\delta(K-B(S,S,T))}\,.
$$

Valuation of caps and floors

- This is equal (up to a change in notional given by the constant factor $1/\widetilde{K}$) to the price of a put on the zero coupon bond struck at $\widetilde{K}$, namely

$$\frac{V(0)}{P_0(0,T)}\,.$$

- Choosing, for now, the underlying process to be lognormal (given by (1)), we conclude that the price of a call on $L(S,T)$ (or caplet) is given by

$$P^{\mathrm{caplet}}(T,K,L_0,\sigma)=\delta P_0(0,T)B^{\mathrm{call}}(S,K,L_0,\sigma), \tag{10}$$

  where L_0 denotes here today's value of the LIBOR forward, namely $L(0,S,T)=L_0(S,T)$.

Valuation of caps and floors

- Since a cap is a basket of caplets, its value is the sum of the values of the constituent caplets:

$$P^{\text{cap}} = \sum_{j=1}^{n} \delta_j B^{\text{call}}(T_{j-1}, K, L_j, \sigma_j) P_0(0, T_j), \tag{11}$$

where δ_j is the day count fraction applying to the accrual period starting at T_{j-1} and ending at T_j, and L_j is the LIBOR forward rate for that period.

- Notice that, in the formula above, the date T_{j-1} has to be adjusted to accurately reflect the expiration date of the option (2 business days before the start of the accrual period).
- Similarly, the value of a floor is

$$P^{\text{floor}} = \sum_{j=1}^{n} \delta_j B^{\text{put}}(T_{j-1}, K, L_j, \sigma_j) P_0(0, T_j). \tag{12}$$

Valuation of caps and floors

- What is an at the money (ATM) cap?
- A key property of an ATM option is that the call and put struck ATM have the same value. We shall first derive a put / call parity relation for caps and floors.
- Let E^{Q_j} denote expected value with respect to the T_j-forward measure Q_{T_j}. Then,

$$
\begin{aligned}
& P^{\text{floor}} - P^{\text{cap}} \\
& \quad = \sum_{j=1}^{n} \delta_j \left(\mathrm{E}^{Q_j} \left[\max \left(K - L_j, 0 \right) \right] - \mathrm{E}^{Q_j} \left[\max \left(L_j - K, 0 \right) \right] \right) P_0(0, T_j) \\
& \quad = \sum_{j=1}^{n} \delta_j \mathrm{E}^{Q_j} \left[K - L_j \right] P_0(0, T_j).
\end{aligned}
$$

Valuation of caps and floors

- Now, the expected value $\mathrm{E}^{\mathrm{Q}_j}\left[L_j\right]$ is the current value of the LIBOR forward $L_0(T_{j-1}, T_j)$.
- We thus arrive at the following put / call parity relation:

$$P^{\text{floor}} - P^{\text{cap}} = K\sum_{j=1}^{n}\delta_j P_0(0, T_j) - \sum_{j=1}^{n}\delta_j L_0(T_{j-1}, T_j) P_0\left(0, T_j\right), \tag{13}$$

which is the present value of the swap receiving K on the quarterly, act/360 basis.

Valuation of caps and floors

- This is an important relation. It implies that:
  - (i) It is natural to think about a floor as a call option, and a cap as a put option.
  - (ii) The underlying asset is the forward starting swap on which both legs pay quarterly, and interest accrues on the act/360 basis.
  - (iii) The coupon dates on the swap coincide with the payment dates on the cap / floor.
  - (iv) The ATM rate is the break-even rate on this swap. This rate is close to but not identical to the break-even rate on the standard semi-annual swap.

Valuation of swaptions

- Consider a swap that settles at T_0 and matures at T, and we let $S(t, T_0, T)$ denote the corresponding (break-even) forward swap rate observed at time $t < T_0$ (in particular, $S_0(T_0, T) = S(0, T_0, T)$).
- We know from Lecture Notes #1 that the forward swap rate is given by

$$S(t, T_0, T) = \frac{\sum_{1 \leq j \leq n_{\mathrm{f}}} \delta_j L_j P(t, T_{\mathrm{val}}, T_j^{\mathrm{f}})}{A(t, T_{\mathrm{val}}, T_0, T)}\,, \tag{14}$$

  where $T_{\mathrm{val}} \leq T_0$ is the valuation date of the swap (its choice has no impact on the value of the rate).
- Here, B_j is the LIBOR / OIS spread, and $A(t, T_{\mathrm{val}}, T_0, T)$ is the forward annuity function:

$$A(t, T_{\mathrm{val}}, T_0, T) = \sum_{1 \leq j \leq n_{\mathrm{c}}} \alpha_j P(t, T_{\mathrm{val}}, T_j^{\mathrm{c}}). \tag{15}$$

Valuation of swaptions

- Using formula (17) of Lecture Notes #1, we write this as

$$\begin{aligned} S(t, T_0, T) &= \frac{\sum_{1 \le j \le n_{\mathrm{f}}} \delta_j L_j P(t, t, T_j^{\mathrm{f}})}{A(t, t, T_0, T)} \\ &= \frac{\sum_{1 \le j \le n_{\mathrm{f}}} \delta_j F_j P(t, t, T_j^{\mathrm{f}})}{A(t, t, T_0, T)} + \frac{\sum_{1 \le j \le n_{\mathrm{f}}} \delta_j B_j P(t, t, T_j^{\mathrm{f}})}{A(t, t, T_0, T)} \\ &= \frac{P(t, t, T_0) - P(t, t, T)}{A(t, t, T_0, T)} + \frac{\sum_{1 \le j \le n_{\mathrm{f}}} \delta_j B_j P(t, t, T_j^{\mathrm{f}})}{A(t, t, T_0, T)} . \end{aligned}$$

- The forward annuity function $A(t, t, T_0, T)$ is the time t present value of an *annuity* paying \$1 on the dates $T_1^{\mathrm{c}}, \ldots, T_{n_{\mathrm{c}}}^{\mathrm{c}}$, as observed at t.
- As in the case of a simple LIBOR forward, the interpretation of (31) is that $S(t, T_0, T)$ is a tradable asset if we use the annuity as numeraire (recall that we are assuming that all the spreads are deterministic!).

Valuation of swaptions

- Indeed, consider the following trade:
  (i) Buy \$1 face value of the zero coupon bond for maturity T_0.
  (ii) Sell \$1 face value of the zero coupon bond for maturity T.
  (iii) Buy a stream of $\delta_j B_j$ face value zero coupon bonds for maturity T_j^{f}, $j = 1, \ldots, n_{\mathrm{f}}$.
- A forward swap rate can thus be modeled as a martingale! Recall that the EMM associated with the annuity function as numeraire is called the *swap measure*.
- A calculation, analogous to the one we did in the case of a caplet, shows that a swaption is equivalent to an option on the forward swap rate, provide we have chosen the annuity function as a numeraire.
- Specifically, the value of a receiver swaption struck at K, expressed in the units of the annuity function, is given by

$$\mathrm{E}^{\mathrm{Q}_{T_0,T}}\left[(K - S(T_0, T))^+\right], \tag{16}$$

where $\mathrm{Q}_{T_0,T}$ denotes the swap (martingale) measure. A receiver (call) option is thus equivalent to a put option on the forward swap rate.

Valuation of swaptions

- Choosing, again, the lognormal process (1), we conclude that today's value of a receiver swaption is thus given by

$$P^{\text{rec}} = A_0(T_0, T) B^{\text{put}}(T_0, K, S_0, \sigma). \tag{17}$$

Similarly the value of a payer swaption is

$$P^{\text{pay}} = A_0(T_0, T) B^{\text{call}}(T_0, K, S_0, \sigma). \tag{18}$$

- Here $A_0(T_0, T) = A(0, 0, T_0, T)$, i.e.

$$A_0(T_0, T) = \sum_{1 \leq j \leq n_c} \alpha_j P_0(0, T_j^c) \tag{19}$$

(all discounting is done to today), and S_0 is today's value of the forward swap rate $S_0(T_0, T)$.

Valuation of swaptions

- The put / call parity relation for swaptions is easy to establish, namely

$$P^{\text{rec}} - P^{\text{pay}} = \text{PV of the swap paying } K \text{ on the semi-annual, 30/360 basis.}$$

- Therefore,
  - (i) It is natural to think about a receiver as a call option, and a payer as a put option.
  - (ii) The ATM rate is the break-even rate on the underlying forward starting swap.

Beyond Black's model

- The basic premise of Black's model, that σ is independent of T, K, and F_0, is not supported by the interest rates markets.
- In fact, option implied volatilities exhibit:
  (i) *Term structure*: At the money volatility depends on the option expiration.
  (ii) *Smile* (or *skew*): For a given expiration, there is a pronounced dependence of implied volatilities on the option strike.
- These phenomena became pronounced in the mid nineties or so and, in order to accurately value and risk manage options portfolios, refinements to Black's model are necessary.
- Modeling term structure of volatility is hard, and not much progress has been made. We will focus on modeling volatility smile.

Local volatility models

- An improvement over Black's model is a class of models called *local volatility models*.
- The idea is that even though the exact nature of volatility (it could be stochastic) is unknown, one can, in principle, use the market prices of options in order to recover the risk neutral probability distribution of the underlying asset.
- To see this, note that

$$\begin{aligned}\frac{d}{dK}(S-K)^{+} &= \frac{d}{dK}\left((S-K)\theta(S-K)\right)\\ &= -\theta(S-K)-(S-K)\delta(S-K)\\ &= -\theta(S-K),\end{aligned}$$

and thus

$$\frac{d^2}{dK^2}(S-K)^{+} = \delta(S-K).$$

Local volatility models

- This implies that

$$\frac{d^2}{dK^2}\,\mathrm{E}^{\mathrm{Q}}[(S_T-K)^+]=\mathrm{E}^{\mathrm{Q}}[\delta(S_T-K)].$$

- Let $g_T(S,S_0)$ denote the terminal probability distribution function of the forward swap rate S. From the equality above we infer that

$$\begin{aligned}\frac{d^2}{dK^2}\,\mathrm{E}^{\mathrm{Q}}[(S_T-K)^+]&=\int\delta(S-K)g_T(S,S_0)dS\\&=g_T(K,S_0).\end{aligned}$$

- Consequently, the terminal probability distribution can (in principle) be computed from the option prices.

Local volatility models

- This, in turn, will allow us to find an effective ("local") specification of the underlying process so that the implied volatilities match the market implied volatilities.
- Such specifications can be justified by developing suitable underlying economic models but, for our purposes, these underlying micro models will be irrelevant, and we will regard a local model as a fit to the observed market.
- Local volatility models are usually specified in the form

$$dF(t) = C(t, F(t))dW(t), \tag{20}$$

where $C(t, F)$ is a certain effective instantaneous volatility.

Local volatility models

- In general, $C(t, F(t))$ is not given in a parametric form. It is, however, a matter of convenience to work with a parametric specification that fits the market data best.
- Popular local volatility models which admit analytic solutions include:
  (i) The normal model.
  (ii) The shifted lognormal model.
  (iii) The CEV model.
- We now briefly discuss the basic features of these models.

Normal model

- The dynamics for the forward rate $F(t)$ in the normal model reads

$$dF(t) = \sigma dW(t), \tag{21}$$

under a suitable choice of numeraire.

- The parameter σ is appropriately called the *normal volatility*.
- This is easy to solve:

$$F(t) = F_0 + \sigma W(t). \tag{22}$$

This solution exhibits one of the main drawbacks of the normal model: with non-zero probability, $F(t)$ may become negative in finite time. Under typical circumstances, this is, however, a relatively unlikely event.

Normal model

- Prices of European calls and puts are now given by:

$$\begin{aligned} P^{\text{call}}(T, K, F_0, \sigma) &= \mathcal{N}(0)\, B_n^{\text{call}}(T, K, F_0, \sigma), \\ P^{\text{put}}(T, K, F_0, \sigma) &= \mathcal{N}(0)\, B_n^{\text{put}}(T, K, F_0, \sigma). \end{aligned} \tag{23}$$

- The functions $B_n^{\text{call}}(T, K, F_0, \sigma)$ and $B_n^{\text{put}}(T, K, F_0, \sigma)$ are given by:

$$\begin{aligned} B_n^{\text{call}}(T, K, F_0, \sigma) &= \sigma\sqrt{T}\Big(d_+ N(d_+) + N'(d_+)\Big), \\ B_n^{\text{put}}(T, K, F_0, \sigma) &= \sigma\sqrt{T}\Big(d_- N(d_-) + N'(d_-)\Big), \end{aligned} \tag{24}$$

where

$$d_\pm = \pm \frac{F_0 - K}{\sigma\sqrt{T}}\ . \tag{25}$$

Normal model

- In order to see it, we write (22) as $F(t) = F_0 + \sigma\sqrt{t}X$, with $X \sim N(0,1)$. In the case of a call option,

$$
\begin{aligned}
\mathrm{E}[(F(T)-K)^+] &= \mathrm{E}[(F_0+\sigma\sqrt{T}X-K)^+] \\
&= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}(\sigma\sqrt{T}X+F_0-K)^+\,e^{-x^2/2}dx \\
&= \frac{1}{\sqrt{2\pi}}\int_{-(F_0-K)/\sigma\sqrt{T}}^{\infty}(F_0-K+\sigma\sqrt{T}x)\,e^{-x^2/2}dx \\
&= \sigma\sqrt{T}\,\frac{1}{\sqrt{2\pi}}\Big(d_+\int_{d_-}^{\infty}e^{-x^2/2}dx+\int_{d_-}^{\infty}xe^{-x^2/2}dx\Big) \\
&= \sigma\sqrt{T}\,\frac{1}{\sqrt{2\pi}}\Big(d_+\int_{-\infty}^{d_+}e^{-x^2/2}dx+e^{-d_+^2/2}dx\Big) \\
&= \sigma\sqrt{T}\Big(d_+N(d_+)+N'(d_+)\Big).
\end{aligned}
$$

- The calculation in the case of a put proceeds along the same lines.

Normal model

- The normal model is (in addition to the lognormal model) an important benchmark in terms of which implied volatilities are quoted.
- In fact, many traders are in the habit of thinking in terms of normal implied volatilities, as the normal model often seems to capture the rates dynamics better than the lognormal (Black's) model.

Shifted lognormal model

- The *shifted lognormal model* (also known as the *displaced diffusion model*) is a diffusion process whose volatility structure is a linear interpolation between the normal and lognormal volatilities.
- Its dynamics reads:

$$dF(t) = (\sigma_1 F(t) + \sigma_0)\, dW(t)\,. \tag{26}$$

  The volatility structure of the shifted lognormal model is given by the values of the parameters σ_1 and σ_0.
- Prices of calls and puts are given by the following valuation formulas:

$$\begin{aligned} P^{\text{call}}(T, K, F_0, \sigma_0, \sigma_1) &= \mathcal{N}(0)\, B_{\text{sln}}^{\text{call}}(T, K, F_0, \sigma_0, \sigma_1), \\ P^{\text{call}}(T, K, F_0, \sigma_0, \sigma_1) &= \mathcal{N}(0)\, B_{\text{sln}}^{\text{put}}(T, K, F_0, \sigma_0, \sigma_1). \end{aligned} \tag{27}$$

Shifted lognormal model

- The functions $B_{\mathrm{sln}}^{\mathrm{call}}(T, K, F_0, \sigma_0, \sigma_1)$ and $B_{\mathrm{sln}}^{\mathrm{put}}(T, K, F_0, \sigma_0, \sigma_1)$ are generalizations of the corresponding functions for the lognormal and normal models:

$$B_{\mathrm{sln}}^{\mathrm{call}}(T, K, F_0, \sigma_0, \sigma_1) = \left(F_0 + \frac{\sigma_0}{\sigma_1}\right) N(d_+) - \left(K + \frac{\sigma_0}{\sigma_1}\right) N(d_-), \tag{28}$$

where

$$d_\pm = \frac{\log \dfrac{\sigma_1 F_0 + \sigma_0}{\sigma_1 K + \sigma_0} \pm \dfrac{1}{2}\, \sigma_1^2 T}{\sigma_1 \sqrt{T}}, \tag{29}$$

and

$$B_{\mathrm{sln}}^{\mathrm{put}}(T, K, F_0, \sigma_0, \sigma_1) = -\left(F_0 + \frac{\sigma_0}{\sigma_1}\right) N(-d_+) + \left(K + \frac{\sigma_0}{\sigma_1}\right) N(-d_-). \tag{30}$$

- The shifted lognormal model is used by some market practitioners as a convenient compromise between the normal and lognormal models. It captures some aspects of the volatility smile.

CEV model

- Another model interpolating between the normal and lognormal models is the *CEV model*[2], whose volatility structure is a power interpolation between the normal and lognormal volatilities.
- Its dynamics is explicitly given by

$$dF(t) = \sigma F(t)^{\beta} \, dW(t),$$

  where $\beta < 1$.
- Note that the exponent β is allowed to be negative. In order for the dynamics to be well defined, we have to prevent $F(t)$ from becoming negative (otherwise $F(t)^{\beta}$ would turn imaginary!).

[2] CEV stands for "constant elasticity of variance".

CEV model

- To this end, we specify a boundary condition at $F = 0$. It can be
  - (i) *Dirichlet* (absorbing): $F|_0 = 0$. Solution exists for all values of β, or
  - (ii) *Neumann* (reflecting): $F'|_0 = 0$. Solution exists for $\frac{1}{2} \leq \beta < 1$.
- Unlike the models discussed above, where the option valuation formulas can be obtained by purely probabilistic methods, the CEV model requires solving a terminal value problem for a partial differential equation, namely the backward Kolmogorov equation:

$$
\begin{aligned}
&\frac{\partial}{\partial t} B(t, f) + \frac{1}{2} \sigma^2 f^{2\beta} \frac{\partial^2}{\partial f^2} B(t, f) = 0, \\
&B(T, f) = \begin{cases} (f - K)^+, & \text{for a call,} \\ (K - f)^+, & \text{for a put,} \end{cases}
\end{aligned}
\tag{31}
$$

This equation has to be supplemented by a boundary condition, Dirichlet of Neumann, at zero f.

CEV model

- Pricing formulas for the CEV model can be obtained in a closed (albeit somewhat complicated) form. For example, in the Dirichlet case the prices of calls and puts are:

$$\begin{aligned} P^{\text{call}}(T, K, F_0, \sigma) &= \mathcal{N}(0)\, B^{\text{call}}_{\text{CEV}}(T, K, F_0, \sigma), \\ P^{\text{put}}(T, K, F_0, \sigma) &= \mathcal{N}(0)\, B^{\text{put}}_{\text{CEV}}(T, K, F_0, \sigma). \end{aligned} \tag{32}$$

- The functions $B^{\text{call}}_{\text{CEV}}(T, K, F_0, \sigma)$ and $B^{\text{put}}_{\text{CEV}}(T, K, F_0, \sigma)$ are the time $t = 0$ solutions to the terminal value problem (31).

CEV model

- They can be expressed in terms of the cumulative function of the non-central χ^2 distribution:
$$\chi^2(x; r, \lambda) = \int_0^x p(y; r, \lambda)\, dy, \qquad (33)$$
whose density is given by a Bessel function [6]:
$$p(x; r, \lambda) = \frac{1}{2}\left(\frac{x}{\lambda}\right)^{(r-2)/4} \exp\left(-\frac{x+\lambda}{2}\right) I_{(r-2)/2}\left(\sqrt{\lambda x}\right). \qquad (34)$$
- Let us also define the quantity:
$$\nu = \frac{1}{2(1-\beta)}, \quad \text{i.e. } \nu \geq \frac{1}{2}. \qquad (35)$$

CEV model

- A tedious (but fun!) computation shows then that the valuation formulas for calls and puts under the CEV model with the Dirichlet boundary condition read:

$$\begin{aligned} B^{\text{call}}_{\text{CEV}}(T, K, F_0, \sigma) = F_0 \Big(1 - \chi^2\Big(\frac{4\nu^2 K^{1/\nu}}{\sigma^2 T};\ 2\nu + 2,\ \frac{4\nu^2 F_0^{1/\nu}}{\sigma^2 T}\Big)\Big) \\ - K\chi^2\Big(\frac{4\nu^2 F_0^{1/\nu}}{\sigma^2 T};\ 2\nu,\ \frac{4\nu^2 K^{1/\nu}}{\sigma^2 T}\Big)\ , \end{aligned} \tag{36}$$

and

$$\begin{aligned} B^{\text{put}}_{\text{CEV}}(T, K, F_0, \sigma) = F_0 \chi^2\Big(\frac{4\nu^2 K^{1/\nu}}{\sigma^2 T};\ 2\nu + 2,\ \frac{4\nu^2 F_0^{1/\nu}}{\sigma^2 T}\Big) \\ - K\Big(1 - \chi^2\Big(\frac{4\nu^2 F_0^{1/\nu}}{\sigma^2 T};\ 2\nu,\ \frac{4\nu^2 K^{1/\nu}}{\sigma^2 T}\Big)\Big)\ , \end{aligned} \tag{37}$$

respectively.

CEV model

- From these formulas one can deduce that the terminal probability density $g_T(F, F_0)$ is given by

$$g_T(F, F_0) = \frac{4\nu F_0 F^{1/\nu-2}}{\sigma^2 T}\, p\left(\frac{4\nu^2 F^{1/\nu}}{\sigma^2 T};\ 2\nu+2,\ \frac{4\nu^2 F_0^{1/\nu}}{\sigma^2 T}\right). \tag{38}$$

- This is the "transition portion" of the process only. Indeed, the total mass of the density $g_T(F, F_0)$ is less than one, meaning that there is a nonzero probability of absorption at zero. Using the series expansion [6]:

$$I_\nu(z) = \sum_{k\geq 0} \frac{1}{k!\,\Gamma(\nu+k+1)} \left(\frac{z}{2}\right)^{2k+\nu}, \tag{39}$$

we readily find that

$$\int_0^\infty g_T(F, F_0)\, dF = 1 - \frac{1}{\Gamma(\nu)}\, \Gamma\left(\nu,\ \frac{2\nu^2 F_0^{1/\nu}}{\sigma^2 T}\right), \tag{40}$$

where

$$\Gamma(\nu,\ x) = \int_x^\infty t^{\nu-1} e^{-t} dt \tag{41}$$

CEV model

- The quantity
$$\frac{1}{\Gamma(\nu)}\,\Gamma\Big(\nu,\ \frac{2\nu^2 F_0^{1/\nu}}{\sigma^2 T}\Big) \tag{42}$$
is the probability of absorption at zero. For example, in the square root process case, i.e. $\nu = 1$, that probability equals $\exp(-\frac{2F_0}{\sigma^2 T})$.
- The total terminal probability is thus the sum of $g_T(F, F_0)$ and the Dirac delta function $\delta(F)$ multiplied by the absorption probability (42).
- Similar valuation formulas hold for the Neumann boundary condition but we will not reproduce them here (see the note [5] for details)

Beyond local volatility models

- The volatility skew models that we have discussed so far improve on Black's models but still fail to accurately reflect the market dynamics.
- One issue is, for example, the "wing effect" exhibited by the implied volatilities of some maturities (especially shorter dated) and tenors which is not captured by these models: the implied volatilities tend to rise for high strikes forming the familiar "smile" shape.
- Among the attempts to move beyond the locality framework are:
  - (i) *Stochastic volatility models*. In this approach, we add a new stochastic factor to the dynamics by assuming that a suitable volatility parameter itself follows a stochastic process.
  - (ii) *Jump diffusion models*. These models use a broader class of stochastic processes (for example, *Levy processes*) to drive the dynamics of the underlying asset. These more general processes allow for discontinuities ("jumps") in the asset dynamics.
- Because of time constraints we shall limit our discussion to an example of approach (i), namely the SABR stochastic volatility model [2].

Dynamics of SABR

- The SABR model is an extension of the CEV model in which the volatility parameter is assumed to follow a stochastic process.
- Its dynamics is explicitly given by:

$$\begin{aligned} dF(t) &= \sigma(t)\, C(F(t))\, dW(t)\,, \\ d\sigma(t) &= \alpha\sigma(t)\, dZ(t)\,. \end{aligned} \tag{43}$$

- Here $F(t)$ is the forward rate process which, depending on context, may denote a LIBOR forward or a forward swap rate[3], and $\sigma(t)$ is the stochastic volatility parameter. The process is driven by two Brownian motions, $W(t)$ and $Z(t)$, with

$$\mathsf{E}\left[dW(t)\, dZ(t)\right] = \rho dt,$$

where the correlation ρ is assumed constant.

[3] The SABR model specification is also used in markets other than interest rate market, and thus $F(t)$ may denote e.g. a crude oil forward.

Dynamics of SABR

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

$$C(F) = F^{\beta}. \tag{44}$$

- Note that we assume that a suitable numeraire has been chosen so that $F(t)$ is a martingale. 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.
- As usual, we supplement the dynamics with the initial condition

$$\begin{aligned} F(0) &= F_0, \\ \sigma(0) &= \sigma_0, \end{aligned} \tag{45}$$

  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 (43) requires a boundary condition at $F = 0$. One usually imposes the absorbing (Dirichlet) boundary condition.

SABR PDE

- As in the case of the CEV model, the analysis of the SABR model requires solving the terminal value problem for the backward Kolmogorov equation associated with the process (43).
- Namely, the valuation function $B = B(t, K, x, \Sigma, y)$ (where x corresponds to the forward F and y corresponds to the volatility parameter σ) is the solution to the following terminal value problem:

$$\begin{aligned} &\frac{\partial}{\partial t} B + \frac{1}{2} y^2 \Big(x^{2\beta} \frac{\partial^2}{\partial x^2} + 2\alpha\rho x^\beta \frac{\partial^2}{\partial x \partial y} + \alpha^2 \frac{\partial^2}{\partial y^2} \Big) B = 0, \\ &B(T, K, x, \Sigma, y) = \begin{cases} (x-K)^+ \ \delta(y-\Sigma), & \text{for a call,} \\ (K-x)^+ \ \delta(y-\Sigma), & \text{for a put.} \end{cases} \end{aligned} \tag{46}$$

- Here, Σ is the *unknown* value of the volatility parameter σ at option expiration T.

SABR PDE

- The price of an option (call or put) is then given by

$$P(T, K, F_0, \sigma) = \mathcal{N}(0) \int_0^\infty B(T, K, F_0, \Sigma, \sigma_0) d\Sigma, \tag{47}$$

  i.e. the unknown terminal value of Σ is integrated out.
- From a numerical perspective, this formula is rather cumbersome: it requires solving the three dimensional PDE (46) and then calculating the integral (47)
- 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 to complex to be of practical use.

SABR implied volatility

- The SABR model can be solved approximately by means of a perturbation expansion in the parameter $\varepsilon = T\alpha^2$, where T is the maturity of the option.
- As it happens, this parameter is typically small and the approximate solution is actually quite accurate.
- 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.
- The parameter α shows a persistent stable term structure as a function of the swaption maturity and the tenor of the underlying swap.
- On a given market snapshot, the highest α's 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).
- Typically, α is a monotone decreasing function of both the option expiry and the underlying swap tenor. A typical range of values of α is $0.2 \lessapprox \alpha \lessapprox 2$.[4]

[4] On 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.

SABR implied volatility

- As already mentioned, there is no known closed form option valuation formula in the SABR model.
- Instead, one takes the following approach. We force the valuation formula to be of the form

$$
\begin{aligned}
P^{\text{call}}(T, K, F_0, \sigma_{\text{n}}) &= \mathcal{N}(0)\,\sigma_{\text{n}}\sqrt{T}\Big(d_+ N(d_+) + N'(d_+)\Big),\\
P^{\text{put}}(T, K, F_0, \sigma_{\text{n}}) &= \mathcal{N}(0)\,\sigma_{\text{n}}\sqrt{T}\Big(d_- N(d_-) + N'(d_-)\Big),\\
d_\pm &= \pm\frac{F_0 - K}{\sigma_{\text{n}}\sqrt{T}}\,,
\end{aligned}
\tag{48}
$$

given by the normal model, with the (normal) implied volatility σ_{n} depending on the SABR model parameters.

SABR implied volatility

- A detailed analysis of the SABR PDE shows that the implied normal volatility is then approximately given by:

$$\begin{aligned}\sigma_{\mathrm{n}}(T, K, F_0, \sigma_0, \alpha, \beta, \rho) = \alpha\, \frac{F_0 - K}{D(\zeta)} \Big\{ 1 + \Big[\frac{2\gamma_2 - \gamma_1^2}{24} \\ \times \Big(\frac{\sigma_0 C(F_{\mathrm{mid}})}{\alpha} \Big)^2 + \frac{\rho\gamma_1}{4}\, \frac{\sigma_0 C(F_{\mathrm{mid}})}{\alpha} + \frac{2 - 3\rho^2}{24} \Big] \varepsilon + \dots \Big\}.\end{aligned} \tag{49}$$

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

$$\begin{aligned}\gamma_1 &= \frac{C'(F_{\mathrm{mid}})}{C(F_{\mathrm{mid}})}, \\ \gamma_2 &= \frac{C''(F_{\mathrm{mid}})}{C(F_{\mathrm{mid}})}.\end{aligned}$$

SABR implied volatility

- The distance function $\delta(\zeta)$ entering the formula above is given by:

$$D(\zeta) = \log\left(\frac{\sqrt{1 - 2\rho\zeta + \zeta^2} + \zeta - \rho}{1 - \rho}\right),$$

where

$$\begin{aligned} \zeta &= \frac{\alpha}{\sigma_0} \int_K^{F_0} \frac{dx}{C(x)} \\ &= \frac{\alpha}{\sigma_0 (1 - \beta)} \left(F_0^{1-\beta} - K^{1-\beta}\right). \end{aligned} \tag{50}$$

- Note that these expressions reduce to the corresponding expressions that we derived for the normal SABR model when $\beta = 0$, $\rho = 0$.

SABR implied volatility

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

$$\begin{aligned}\sigma_{\ln}(T,K,F_0,\sigma_0,\alpha,\beta,\rho) = \alpha\,\frac{\log(F_0/K)}{D(\zeta)}\Big\{1+\Big[\frac{2\gamma_2-\gamma_1^2+1/F_{\mathrm{mid}}^2}{24}\\ \times\Big(\frac{\sigma_0 C(F_{\mathrm{mid}})}{\alpha}\Big)^2+\frac{\rho\gamma_1}{4}\,\frac{\sigma_0 C(F_{\mathrm{mid}})}{\alpha}+\frac{2-3\rho^2}{24}\Big]\varepsilon+\dots\Big\}.\end{aligned} \tag{51}$$

Arbitrage in the SABR implied vol approximation

- The explicit implied volatility given by formulas (49) or (51) 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 $\alpha^2 T \ll 1$.
- The unstated argument 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 in the SABR implied vol approximation

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

$$
\begin{aligned}
P^{\text{call}}(T,K) &= \int_{-\infty}^{\infty} (F-K)^{+}\, g_T(F,K)\, dF,\\
P^{\text{put}}(T,K) &= \int_{-\infty}^{\infty} (K-F)^{+}\, g_T(F,K)\, dF,
\end{aligned}
\tag{52}
$$

  where $g_T(F,K)$ is the risk-neutral probability density at the exercise date (including the delta function from the Dirichlet boundary condition).
- As we saw earlier,

$$
\begin{aligned}
\frac{\partial^2}{\partial K^2} P^{\text{call}}(T,K) &= \frac{\partial^2}{\partial K^2} P^{\text{put}}(T,K)\\
&= g_T(F,K)\\
&\geq 0,
\end{aligned}
\tag{53}
$$

  for all K.

Arbitrage in the SABR implied vol approximation

- So the explicit implied volatility formula (49) can represent an arbitrage free model only if

$$\begin{aligned}\frac{\partial^2}{\partial K^2}\,P^{\text{call}}(T,K,F_0,\sigma_{\mathrm{n}}(K,\ldots)) &= \frac{\partial^2}{\partial K^2}\,P^{\text{put}}(T,K,F_0,\sigma_{\mathrm{n}}(K,\ldots))\\ &\geq 0.\end{aligned} \tag{54}$$

- 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. 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 $\sigma_{\mathrm{n}}(K,\ldots))$ which violates the arbitrage free constraint in (53) for a range of values of K.
- This issue is resolved in [4] where a somewhat more refined analysis of the model is presented.

SABR and negative rates

- In the aftermath of the last financial crisis, central banks (particularly in the US and EU) implemented the policy of *quantitative easing* in order to provide liquidity to the markets.
- As a result, some of the rates in the EU and Japanese markets became negative.
- The original version of the SABR model requires rates to be positive. In order to adapt it to the reality of negative rates, market practitioners extended it to the *shifted SABR model*.
- Its dynamics is given by:

$$\begin{aligned} dF(t) &= \sigma(t)\,(F(t)+\theta)^{\beta}\,dW(t)\,, \\ d\sigma(t) &= \alpha\sigma(t)\,dZ(t)\,, \end{aligned} \tag{55}$$

  where $\theta > 0$ is a shift parameter.
- Note that the asymptotic formulas (49) and (51) for the implied volatility apply to the shifted SABR model.

Calibration of SABR

- For each option maturity and underlying we have to specify 4 model parameters: $\sigma_0, \alpha, \beta, \rho$ (and the shift parameter θ in the markets with negative rates).
- In order to do it we need market implied volatilities for several different strikes.
- Given this, the calibration poses no problem: one can use any standard NLS optimization algorithm.
- Calibration results show interesting term structure of the model parameters as functions of the maturity and underlying.
- As we already mentioned, typical is the shape of the parameter α which start out high for short dated options and then declines monotonically as the option expiration increases. This reflects the presence of jumps in the short end of the curve.

Calibration of SABR

- 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 one of these parameters.
- The first approach is to fix $\beta \leq 1$, say $\beta = 0.5$, and calibrate σ_0, α, ρ. This choice of calibration methodology works quite well under "normal" conditions.
- In times of distress, such as 2008 - 2009, the height of the last financial crisis, 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.
- The second approach is to fix $\rho = 0$, and calibrate σ_0, α, β.

Building volatility cube

- We can organize the implied swaption volatilities by:
  (i) Option expiration.
  (ii) Tenor of the underlying swap.
  (iii) Strike on the option.
- This three dimensional object is referred to as the *volatility cube*.
- Swaption market is relatively liquid but, on a given day, only a limited number of structures trade.
- Out of over a hundred elements of the volatility matrix, the market provides information for certain benchmark maturities (1 month, 3 months, 6 months, 1 year, ...), underlyings (1 year, 2 years, 5 years, ...), and strikes (ATM, $\pm$50 bp, ...) only.
- The process of volatility cube construction requires performing intelligent interpolations for the remaining expiries and tenors in a way that does not allow for arbitrage.

Building volatility cube

- The market quotes swaption volatilities for certain standard maturities and underlyings. Table 4 contains a snapshot of the matrix of lognormal (Black) at the money swaption volatilities for the standard expirations and underlyings.

	1Y	2Y	3Y	4Y	5Y	7Y	10Y	15Y	20Y	30Y
1M	71.7%	68.2%	61.8%	56.0%	50.6%	49.7%	47.3%	44.4%	43.4%	43.0%
3M	73.6%	68.8%	61.1%	55.9%	51.5%	49.6%	47.2%	44.0%	42.8%	42.8%
6M	72.2%	70.0%	60.0%	55.6%	51.9%	49.0%	46.1%	42.9%	42.1%	41.9%
1Y	73.8%	65.5%	57.1%	52.9%	49.7%	46.6%	43.8%	40.8%	40.3%	40.0%
2Y	73.7%	59.0%	51.4%	47.2%	45.4%	42.0%	39.4%	37.0%	36.6%	36.4%
3Y	57.8%	49.0%	44.3%	41.8%	40.4%	37.9%	36.0%	34.3%	34.1%	33.4%
4Y	46.0%	41.6%	39.2%	37.8%	36.8%	35.1%	33.9%	32.7%	32.4%	32.1%
5Y	39.7%	37.8%	36.6%	35.6%	34.8%	33.5%	32.6%	31.7%	31.6%	31.3%
7Y	34.4%	33.4%	32.4%	31.7%	31.2%	30.7%	30.7%	30.0%	29.7%	29.2%
10Y	29.8%	29.4%	29.1%	28.8%	28.7%	28.8%	29.2%	28.2%	27.5%	26.8%

Table: 4. Swaption ATM lognormal volatilities

Building volatility cube

- Alternatively, Table 5 contains the matrix of normal at the money swaption volatilities expressed, for ease of readability, in basis points.

	1Y	2Y	3Y	4Y	5Y	7Y	10Y	15Y	20Y	30Y
1M	48	48	51	58	65	85	101	111	114	116
3M	51	49	53	61	69	88	102	111	113	116
6M	51	52	56	66	75	91	103	110	112	115
1Y	52	56	63	74	82	94	104	108	110	111
2Y	72	76	83	88	94	99	103	104	104	104
3Y	91	93	95	98	100	101	103	102	101	99
4Y	101	101	101	102	102	102	103	100	98	97
5Y	105	104	104	104	104	103	102	99	97	95
7Y	103	102	101	100	100	99	98	94	92	89
10Y	98	97	96	95	94	93	92	87	84	81

Table: 5. Swaption ATM normal volatilities (in basis points)

Building volatility cube

- The matrix of at the money volatilities should be accompanied by the matrix of forward swap rates, calculated from the rate market snapshot. These are the at the money strikes for the corresponding swaptions. This is illustrated by Table 6.

	1Y	2Y	3Y	4Y	5Y	7Y	10Y	15Y	20Y	30Y
1M	0.67%	0.70%	0.82%	1.04%	1.29%	1.72%	2.13%	2.49%	2.62%	2.71%
3M	0.69%	0.72%	0.87%	1.10%	1.35%	1.77%	2.18%	2.52%	2.64%	2.72%
6M	0.71%	0.76%	0.94%	1.19%	1.45%	1.86%	2.25%	2.57%	2.67%	2.75%
1Y	0.72%	0.87%	1.13%	1.41%	1.66%	2.03%	2.38%	2.66%	2.74%	2.80%
2Y	1.02%	1.33%	1.64%	1.90%	2.10%	2.39%	2.66%	2.84%	2.88%	2.90%
3Y	1.64%	1.96%	2.21%	2.39%	2.52%	2.72%	2.91%	3.00%	3.00%	3.00%
4Y	2.27%	2.50%	2.64%	2.75%	2.84%	2.98%	3.09%	3.11%	3.09%	3.07%
5Y	2.72%	2.84%	2.92%	2.99%	3.06%	3.15%	3.21%	3.18%	3.14%	3.11%
7Y	3.08%	3.16%	3.21%	3.26%	3.29%	3.31%	3.29%	3.22%	3.17%	3.13%
10Y	3.40%	3.41%	3.41%	3.40%	3.38%	3.33%	3.26%	3.18%	3.14%	3.11%

Table: 6. Forward swap rates (ATM strikes)

Stripping cap volatility

- A cap is a basket of options of different maturities and different moneynesses.
- For simplicity, the market quotes cap / floor prices in terms of a single number, the *flat volatility*. This is the single volatility which, when substituted into the valuation formula (for all caplets / floorlets!), reproduces the correct price of the instrument.
- Clearly, flat volatility is a dubious concept: since a single caplet may be part of different caps it gets assigned different flat volatilities.
- The process of constructing actual implied caplet volatility from market quotes is called *stripping* cap volatility.
- The result is a sequence of ATM caplet volatilities for maturities ranging from one day to, say, 30 years. Convenient benchmarks are 3 months, 6 months, 9 months, ... The market data usually include Eurodollar options and OTC caps and floors.

Stripping cap volatility

- There are various methods of stripping at the money cap volatility. Among them we mention:
  - (i) *Bootstrap*. One starts at the short end and moves further trying to match the prices of Eurodollar options and spot starting caps / floors. This method tends to produce a jagged shape of the volatility curve.
  - (ii) *Optimization*. This method produces a smooth shape of the cap volatility curve but is somewhat more involved. We use a two step approach: in the first step fit the caplet volatilities to the *hump function*:

$$H(t) = (\alpha + \beta t)e^{-\lambda t} + \mu. \tag{56}$$

- Generally, the hump function gives a qualitatively correct shape of the cap volatility. Quantitatively, the fit is insufficient for accurate pricing and should be refined. A good approach is to use cubic splines.
- Once α, β, λ, and μ have been calibrated, we use cubic splines in a way similar to the method explained in Lecture Notes #1 in order to nail down the details of the caplet volatility curve.

Volatility cube

- The third dimension of the volatility cube is the strike dependence of volatility.
- It can be conveniently characterized in terms of the parameters of the smile model such as the parameters α, β, and ρ of the SABR model.
- We can conveniently organize these parameters in matrices of the same dimensions as the at the money volatility matrices.

References

- Gatheral, J.: *The Volatility Surface: A Practitioner's Guide*, Wiley (2006).
- 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* (2013).
- Lesniewski, A.: Notes on the CEV model, NYU working paper (2009).
- Whittaker, E. T., and Watson, G. N.: *A Course of Modern Analysis*, Cambridge University Press (1996).