Lognormal vs Normal Volatilities and Sensitivities in Practice

Georgi Dimitroff
georgi.dimitroff@gmail.com

Christian Fries
email@christian-fries.de

Mark Lichtner
lichtner@gmx.net

Niklas Rodi
niklas.rodi@gmail.com

March 19, 2016

Abstract

The advent of close to zero or even negative rates in major currencies has made the traditional lognormal Black-Scholes-Merton volatility as a representation of option prices in the interest rate market obsolete. Recently more and more cap/floor and even swaption prices in major currencies are violating the upper no-arbitrage bound implied by the Black-Scholes-Merton model. The corresponding lognormal volatilities fail to exist and, thus, cannot be used as inputs for trading and risk systems. Consequently many market participants have resorted to either a normal or displaced lognormal volatility market data representation.

Altering the quoting convention and representing option prices in displaced lognormal or normal rather than traditional lognormal volatilities may look like a technical detail. However, depending on the set-up this may have a direct impact on the definition of the risk factor coordinate system. In this case traders and risk managers do not only need to understand the levels of the new types of volatilities but also need to rebuild their intuition for the associated risk sensitivities.

In this paper we present approximate and closed-form formulas to transform lognormal volatilities and sensitivities into their normal or displaced lognormal counterparts and vice versa. These results are crucial for understanding and validating the potentially severe impact on sensitivities and (delta) hedge ratios that may appear when changing the volatility market data representation.

Contents

1 Introduction **2**

2 Definitions **3**

3 Relationship between (displaced) lognormal and normal implied volatilities **5**
- 3.1 Approximate volatility transformations . . . 5
- 3.2 Exact at-the-money volatility transformations . . . 7
- 3.3 Excursion1: Choosing the displacement to quote the smile . . . 8
- 3.4 Excursion2: Transforming volatilities for highly correlated underlyings . . . 9

4 Relationship between (displaced) lognormal and normal Greeks **10**
- 4.1 Set-up . . . 10
- 4.2 Invariance of prices when changing the volatility representation . . . 11
- 4.3 The delta and vega risk factor coordinate systems . . . 11
- 4.4 Normal to lognormal linearisation . . . 13
- 4.5 Delta and vega transformations . . . 15
  - 4.5.1 The volatility cube case . . . 16
  - 4.5.2 The at-the-money volatility case . . . 18

5 Conclusion **19**

1 Introduction

Traditionally swaption and cap/floor prices are quoted in terms of their implied (lognormal) Black-Scholes-Merton volatilities which are used as input for the trading and risk systems.[1] The evolution of close to zero or even negative (forward) LIBOR and swaprates in major currencies has made this representation to become obsolete as more and more option prices violate the no-arbitrage bounds implied by a lognormal model.[2] Consequently, the corresponding buckets disappeared from the quoted volatility cube. This has forced brokers, banks and other market participants to move from a lognormal to either a displaced lognormal or normal volatility representation.

Altering the type of input volatility seems to be just a trivial change in the quotation of market data. As discussed in section 4.2, it has per se no effect on the prices of the interest rate products[3] and the risk of the products remains the same as well. However, what may change is the *representation* of the risk! As (ATM) vegas are usually calculated by bumping the corresponding quoted (ATM) volatilities, a change in the definition of the input volatilities directly influences the vegas. Furthermore, as discussed in section 4.3, if the (ATM) volatilities are part of the vega risk factors then altering their representation goes hand in hand with a change of the whole risk factor coordinate system. In this case the volatility quoting convention has an impact on the deltas as well - at least for volatility sensitive product. Furthermore and maybe even more strikingly, while having no effect on the vega hedge ratios such a change in the risk factor coordinate system affects the delta hedge ratios of every volatility sensitive product - and this to a degree that some hedge ratios may even change their sign.

The formulas presented in this paper can be used as a consistency check when changing the volatility input type and as means to adapt the delta and vega risk limits in a sensible way.

[1] Note that we use the terms Black-Scholes-Merton volatility and Black-76 volatility interchangeably throughout this paper.
[2] See section 2 and section 3.2, especially Remark 3.3, for the arbitrage bounds of a (displaced) lognormal model.
[3] Note however the subtle limitations to this view discussed in section 4.2.

While many of the volatility transformation results were already derived in previous works cited in the bibliography - some formulas like the (rather poor) approximation (3.3) of the normal volatility as the product of forward and lognormal volatility can indeed be considered to be "folklore" - this paper gives a thorough summary and clear derivation of the most important ones. Our results on the transformation of lognormal sensitivities into normal or displaced lognormal ones, especially for the non-ATM case, are new insights to the best of our knowledge.

The outline of this paper is as follows. Section 2 sets the notational frame for this study. Section 3 provides approximate and exact formulas to convert (displaced) lognormal volatilities into normal volatilities and vice versa. Finally, section 4 discusses the transformation between normal and lognormal and between displaced lognormal and lognormal deltas and vegas.

2 Definitions

Although we take the case of a swaption as our favoured application, most results presented here are valid for general classes of European options and also apply for example to caplets and floorlets. In what follows the underlying is a forward swaprate $F(t)|_{t\in[0,T]}$, where 0 is today, T is the option maturity and t denotes a dynamic time variable.

A swap exchanges floating LIBOR rates against a fix rate. To be precise, let $T \leq T_0^{fix}$, $\ldots$, $T_{N_{fix}}^{fix}$ and $T \leq T_0^{float}, \ldots, T_{N_{float}}^{float}$ be the tenor structures of the fix and the floating leg, respectively. Let $L_i(t)$ denote the forward rate at time t of a LIBOR starting in T_i^{float} and payment in T_{i+1}^{float} with corresponding daycount fraction τ_i^{float}. The daycount fraction of the fix rate period $[T_j^{fix}, T_{j+1}^{fix}]$ is denoted by τ_j^{fix}.

The forward swaprate $F(t)$ is then defined as

$$F(t) := \frac{PV_{float}(t)}{A(t)},$$

where

$$PV_{float}(t) := \sum_{i=0}^{N_{float}-1} L_i(t) \cdot \tau_i^{float} \cdot P_d(t, T_{i+1}^{float})$$

is the present value of the floating leg and

$$A(t) := \sum_{j=0}^{N_{fix}-1} \tau_j^{fix} \cdot P_d(t, T_{j+1}^{fix})$$

is the annuity of the fix leg. P_d denotes the discounting factor associated with the swap. Since market swaptions typically reference collateralized swaps, P_d corresponds to the collateral rate, usually EONIA in the EUR market or the fed fund rate in the USD market.

The traditional text book model for quoting market prices of plain vanilla swaptions is the (lognormal) Black-Scholes-Merton model. Here the underlying forward swaprate F_{LN} is modelled as a geometric Brownian motion:

$$F_{LN}(t) = F(0) + \int_0^t F_{LN}(u)\sigma_{LN}\, \mathrm{d}B_u,$$

where σ_{LN} is the lognormal volatility and B_u denotes a Brownian motion in the annuity measure. Obviously, in this model zero is unattainable and all future forward swaprates have the same sign as $F(0)$.

In the displaced/shifted Black-Scholes-Merton model with displacement $d > 0$ the underlying forward swaprate F^d_{LN} is modeled as a displaced geometric Brownian motion:

$$F^d_{LN}(t) = F(0) + \int_0^t (F_{LN}(u) + d)\, \sigma^d_{LN}\, \mathrm{d}B_u,$$

which means that the displaced process $F^d_{LN} + d$ follows a geometric Brownian motion. σ^d_{LN} is the displaced lognormal volatility. Trivially, the classic Black-Scholes-Merton model can be considered as a special case of the displaced Black-Scholes-Merton model by setting the displacement d equal to zero. In particular $\sigma^d_{LN} = \sigma_{LN}$ for $d = 0$.[4]

In the Bachelier normal model the underlying forward swaprate F_N is modelled as a scaled Brownian motion:

$$F_N(t) = F(0) + \sigma_N B_t.$$

where σ_N is the normal (Bachelier) volatility.

In the following we denote today's forward as F instead of $F(0)$ and today's annuity as A instead of $A(0)$ to keep the notation simple. The price of a swaption with strike K and maturity T is given by either:

1. The Black-Scholes-Merton formula for a call option (payer swaption)

$$\begin{aligned} &C_{LN}(A, F, K, T, \sigma_{LN}) \\ &\quad = A\left[F \cdot \Phi\left(\frac{\log\frac{F}{K} + \frac{1}{2}\sigma^2_{LN}T}{\sigma_{LN}\sqrt{T}}\right) - K \cdot \Phi\left(\frac{\log\frac{F}{K} - \frac{1}{2}\sigma^2_{LN}T}{\sigma_{LN}\sqrt{T}}\right)\right] \end{aligned} \tag{2.1}$$

and put option (receiver swaption)

$$\begin{aligned} &P_{LN}(A, F, K, T, \sigma_{LN}) \\ &\quad = A\left[-F \cdot \Phi\left(-\frac{\log\frac{F}{K} + \frac{1}{2}\sigma^2_{LN}T}{\sigma_{LN}\sqrt{T}}\right) + K \cdot \Phi\left(-\frac{\log\frac{F}{K} - \frac{1}{2}\sigma^2_{LN}T}{\sigma_{LN}\sqrt{T}}\right)\right]. \end{aligned} \tag{2.2}$$

Obviously (2.1) and (2.2) are only defined for $\frac{F}{K} > 0$.

2. The displaced Black-Scholes-Merton formula for a call option

$$C^d_{LN}(A, F, K, T, \sigma^d_{LN}) = C_{LN}(A, F + d, K + d, T, \sigma^d_{LN}) \tag{2.3}$$

and put option

$$P^d_{LN}(A, F, K, T, \sigma^d_{LN}) = P_{LN}(A, F + d, K + d, T, \sigma^d_{LN}). \tag{2.4}$$

3. The Bachelier formula for a call option

$$\begin{aligned} &C_N(A, F, K, T, \sigma_N) \\ &\quad = A\left[(F - K) \cdot \Phi\left(\frac{F - K}{\sigma_N\sqrt{T}}\right) + \sigma_N\sqrt{T} \cdot \phi\left(\frac{F - K}{\sigma_N\sqrt{T}}\right)\right] \end{aligned} \tag{2.5}$$

[4] For notational convenience we often do not consider the lognormal model on its own but rather as a special case of the displaced lognormal model. To emphasize these cases we here put the word "displaced" in brackets.

and put option

$$P_N(A,F,K,T,\sigma_N) \\ = A\left[(K-F)\cdot\Phi\left(\frac{K-F}{\sigma_N\sqrt{T}}\right)+\sigma_N\sqrt{T}\cdot\phi\left(\frac{F-K}{\sigma_N\sqrt{T}}\right)\right], \tag{2.6}$$

where $\Phi(x) := \int_{-\infty}^{x}\phi(u)du$ and $\phi(x) := \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$ are the standard normal probability distribution and density function, respectively.

These formulas are used for quoting purposes, meaning that given the current forward swaprate F and the market price of a swaption one inverts (2.1) - (2.6) to obtain the implied (displaced) lognormal or normal volatility σ_{LN}, σ_{LN}^d or σ_N, respectively. The mappings $\sigma_{LN}\mapsto C_{LN}(A,F,K,T,\sigma_{LN})$, $\sigma_{LN}^d\mapsto C_{LN}^d(A,F,K,T,\sigma_{LN}^d)$ and $\sigma_N\mapsto C_N(A,F,K,T,\sigma_N)$ are all strictly increasing and hence invertible on their respective image space.[5] However, while it is always possible to match an arbitrary large swaption price by choosing a sufficiently large normal volatility, this does not work in general using the (displaced) Black-Scholes-Merton formula. It is easy to see from (2.1) - (2.4) that the supremum of the price of a payer or receiver swaption in a (displaced) lognormal model is $A\,(F+d)$ and $A\,(K+d)$, respectively.[6] So if the displacement is not chosen large enough it may happen that the swaption prices are too big to be expressible in terms of a (displaced) lognormal volatility. This is discussed in more detail in section 3.2, see especially Remark 3.3.

3 Relationship between (displaced) lognormal and normal implied volatilities

3.1 Approximate volatility transformations

For the general non-ATM case ($K \neq F$) no closed-form formula exists to transform (displaced) lognormal volatilities into normal volatilities and vice versa. Hence one has to either solve the equation

$$C_{LN}^d(A,F,K,T,\sigma_{LN}^d) = C_N(A,F,K,T,\sigma_N)$$

for σ_{LN}^d or σ_N numerically or to resort to approximations.[7] Following the second approach, [2] has proven that

$$\sigma_N = \sigma_{LN}\cdot F\cdot\frac{\frac{K}{F}-1}{\log(\frac{K}{F})}\left[1-\frac{\log\left(\frac{K-F}{\sqrt{KF\cdot\log(\frac{K}{F})}}\right)}{\log^2\frac{F}{K}}\sigma_{LN}^2T\right]+O(T^2\log T).$$

[5] Though inverting the formulas may impose numerical challenges as described in [3].

[6] This can also be shown in a different way. For a payer swaption (call option) C in any displaced lognormal model one has

$$C = A\cdot\mathbb{E}^{Q_A}\left[\left|F_{LN}^d(T)-K\right|^+\right] = A\cdot\mathbb{E}^{Q_A}\left[\left|(F_{LN}^d(T)+d)-(K+d)\right|^+\right] \\ \leq A\cdot\mathbb{E}^{Q_A}\left[\left|F_{LN}^d(T)+d\right|^+\right] = A\cdot\mathbb{E}^{Q_A}\left[F_{LN}^d(T)+d\right] = A(F+d)$$

as long as $K+d\geq 0$. Similarly for a receiver swaption (put option) P in a displaced lognormal model one has

$$P = A\cdot\mathbb{E}^{Q_A}\left[\left|(K+d)-(F_{LN}^d(T)+d)\right|^+\right] \leq A\cdot\mathbb{E}^{Q_A}\left[|(K+d)|^+\right] = A(K+d) \tag{2.7}$$

as long as $K+d\geq 0$. In fact these bounds hold for any model where the distribution of the underlying F is a shift d of a positive distribution.

[7] C_{LN}^d and C_N are defined by (2.3) and (2.5), respectively.

In particular this implies that

$$\sigma_N = \sigma_{LN} \cdot F \cdot \frac{\frac{K}{F} - 1}{\log(\frac{K}{F})} \quad \text{for } T \to 0 \tag{3.1}$$

This formula can be linearized in the smile direction by expanding around $\frac{K}{F} = 1$. In first order this yields[8]

$$\sigma_N \approx \sigma_{LN} \frac{F + K}{2} \text{ for } T \to 0 \text{ and } \frac{K}{F} \approx 1. \tag{3.2}$$

Taking the ATM limit $K \to F$ one gets the well-known and popular short maturity ATM approximation

$$\sigma_N = \sigma_{LN} \cdot F \text{ for } T \to 0, \tag{3.3}$$

which was originally proven in [4]. This result can be extended to the displaced case

$$\sigma_N = \sigma_{LN}^d \, (F + d) \text{ for } T \to 0 \tag{3.4}$$

using the following proposition

Proposition 3.1. *For an ATM swaption (i.e. $F = K$) the (displaced) lognormal and normal volatilities satisfy the following condition*

$$0 \le \sigma_{LN}^d - \frac{\sigma_N}{F + d} \le \frac{T}{24} {\sigma_{LN}^d}^3.$$

Proof: We follow the proof in [4] and only trivially add a displacement d[9]. Given the market price C^{ATM} of an ATM swaption with maturity T the implied normal ATM volatility σ_N is given by

$$\frac{C^{ATM}}{A} = \sigma_N \sqrt{\frac{T}{2\pi}}$$

and for the implied (displaced) lognormal ATM volatility we have

$$\frac{C^{ATM}}{A} = (F + d) \cdot \left(\Phi\left(\frac{\sigma_{LN}^d \sqrt{T}}{2}\right) - \Phi\left(-\frac{\sigma_{LN}^d \sqrt{T}}{2}\right) \right)$$

This implies that

$$\frac{\sigma_N}{F + d} = \sqrt{\frac{2\pi}{T}} \cdot \left(\Phi\left(\frac{\sigma_{LN}^d \sqrt{T}}{2}\right) - \Phi\left(-\frac{\sigma_{LN}^d \sqrt{T}}{2}\right) \right)$$

and, hence, using $1 - x \le e^{-x}$ for $x \ge 0$

$$\begin{aligned}
\sigma_{LN}^d - \frac{\sigma_N}{F + d} &= \sigma_{LN}^d - \sqrt{\frac{2\pi}{T}} \cdot \left(\Phi\left(\frac{\sigma_{LN}^d \sqrt{T}}{2}\right) - \Phi\left(-\frac{\sigma_{LN}^d \sqrt{T}}{2}\right) \right) \\
&= \sqrt{\frac{1}{T}} \cdot \left(\sigma_{LN}^d \sqrt{T} - \int_{-\frac{1}{2}\sigma_{LN}^d \sqrt{T}}^{\frac{1}{2}\sigma_{LN}^d \sqrt{T}} e^{-\frac{u^2}{2}} \, \mathrm{d}u \right) \\
&= \sqrt{\frac{1}{T}} \cdot \int_{-\frac{1}{2}\sigma_{LN}^d \sqrt{T}}^{\frac{1}{2}\sigma_{LN}^d \sqrt{T}} \left(1 - e^{-\frac{u^2}{2}}\right) \mathrm{d}u \le \sqrt{\frac{1}{T}} \cdot \int_{-\frac{1}{2}\sigma_{LN}^d \sqrt{T}}^{\frac{1}{2}\sigma_{LN}^d \sqrt{T}} \frac{1}{2} u^2 \, \mathrm{d}u = \frac{T}{24} {\sigma_{LN}^d}^3
\end{aligned}$$

□

[8] It is $\frac{x-1}{\log(x)} = 1 + \frac{1}{2} \cdot (x - 1) + O(x^2)$ and with $x = \frac{K}{F}$ we find $\frac{\frac{K}{F}-1}{\log(\frac{K}{F})} = 1 + \frac{1}{2} \cdot (\frac{K}{F} - 1) + O\left((\frac{K}{F})^2\right)$.

[9] There is a factor $\frac{1}{2}$ missing in the original proof in [4].

3.2 Exact at-the-money volatility transformations

Not surprisingly formulas (3.1) - (3.4) do not work well for medium or longer maturities T (e.g. $T \geq 5$). When using these simple formulas to transform sensitivities from lognormal to normal, the results were often found to be very poor. Luckily, in the ATM case there is an exact closed-form relationship between normal and lognormal volatilities [1, 2]:

Proposition 3.2. *Assume an ATM swaption with implied lognormal volatility σ_{LN}. Then the corresponding implied normal volatility is given by*

$$\sigma_N = \sqrt{\frac{2\pi}{T}} \cdot F \cdot \left(2\Phi\left(\frac{\sigma_{LN}\sqrt{T}}{2} \right) - 1 \right) = \sqrt{\frac{2\pi}{T}} \cdot F \cdot \operatorname{erf}\left(\frac{\sigma_{LN}\sqrt{T}}{2\sqrt{2}} \right). \tag{3.5}$$

Assume an ATM swaption with implied displaced lognormal volatility σ^d_{LN}. Then the corresponding implied normal volatility is given by

$$\sigma_N = \sqrt{\frac{2\pi}{T}} \cdot (F+d) \cdot \left(2\Phi\left(\frac{\sigma^d_{LN}\sqrt{T}}{2} \right) - 1 \right) = \sqrt{\frac{2\pi}{T}} \cdot (F+d) \cdot \operatorname{erf}\left(\frac{\sigma^d_{LN}\sqrt{T}}{2\sqrt{2}} \right). \tag{3.6}$$

Proof: Evaluating the equality

$$C_N(A, F, K, T, \sigma_N) = C^d_{LN}(A, F, K, T, \sigma^d_{LN})$$

for the ATM case $F = K$ yields

$$\begin{aligned} \sigma_N\sqrt{T} \cdot \phi(0) &= (F+d)\left[\Phi\left(\frac{\sigma^d_{LN}\sqrt{T}}{2} \right) - \Phi\left(-\frac{\sigma^d_{LN}\sqrt{T}}{2} \right) \right] \\ \Leftrightarrow \sigma_N &= \sqrt{\frac{2\pi}{T}}\,(F+d)\left(2\Phi\left(\frac{\sigma^d_{LN}\sqrt{T}}{2} \right) - 1 \right) \\ &= \sqrt{\frac{2\pi}{T}}\,(F+d) \cdot \frac{2}{\sqrt{2\pi}} \int_0^{\frac{1}{2}\sigma^d_{LN}\sqrt{T}} e^{-\frac{u^2}{2}}\,du. \end{aligned}$$

Using the definition of the error function

$$\operatorname{erf}(x) := \frac{2}{\sqrt{\pi}} \int_0^x e^{-u^2}\,du.$$

proofs the proposition.

As the error function is a strictly increasing function converging to 1 for $x \to \infty$, we get the following:

Remark 3.3. Assume an ATM swaption with implied normal volatility σ_N. Then the swaption price can be expressed in terms of a displaced lognormal volatility if and only if the displacement satisfies the condition

$$F + d > \sqrt{\frac{T}{2\pi}}\sigma_N \tag{3.7}$$

In this case formula (3.6) can be inverted to obtain

$$\sigma^d_{LN} = \frac{2\sqrt{2}}{\sqrt{T}} \cdot \operatorname{erf}^{-1}\left(\frac{\sigma_N}{F+d}\sqrt{\frac{T}{2\pi}} \right). \tag{3.8}$$

In particular, this means that an ATM swaption price can not be expressed in terms of a classical Black-Scholes-Merton lognormal volatility if and only if $F \leq \sqrt{\frac{T}{2\pi}}\sigma_N$.

Remark 3.4. Assume an ATM swaption with implied displaced lognormal volatility σ_{LN}^{d}. Let $\tilde{d}$ be another displacement satisfying[10]

$$F + \tilde{d} > (F + d) \cdot erf\left(\frac{\sigma_{LN}^{d}\sqrt{T}}{2\sqrt{2}}\right)$$

Then the displaced lognormal volatility $\sigma_{LN}^{\tilde{d}}$ exists and is given by

$$\sigma_{LN}^{\tilde{d}} = \frac{2\sqrt{2}}{\sqrt{T}} \cdot \mathrm{erf}^{-1}\left(\frac{F+d}{F+\tilde{d}} \cdot \mathrm{erf}\left(\frac{\sigma_{LN}^{d}\sqrt{T}}{2\sqrt{2}}\right)\right) \tag{3.9}$$

Equation (3.7) gives a simple requirement on the displacement in terms of maturity, forward and normal volatility that needs to be satisfied in order for the (displaced) lognormal volatility to exist. The lower the interest rate F, the larger the maturity T or the larger the normal volatility σ_N the larger one has to choose the displacement d. This explains why in the current low rates environment in the EUR market a displacement of zero was too small for certain long dated swaptions and the corresponding lognormal volatility quotes have disappeared. A generalization of (3.7), taking into account non-ATM strikes, is given by formula (3.10).

3.3 Excursion1: Choosing the displacement to quote the smile

Two popular models used to define or interpolate the volatility smile are the SABR and the Heston model. However, both are only defined for positive interest rates or - in their displaced versions - for interest rates above a certain threshold.[11] So in practice the question arises how to choose a single displacement d to be able to represent the full smile.

Assume that the market quotes payer and/or receiver swaption prices for strikes $K_i = F + \Delta_i$, where F is the forward swaprate and Δ_i the offset. As discussed in section 2 the upper bounds of a payer and receiver swaption with strike K in a displaced lognormal model are $A\,(F+d)$ and $A\,(K+d)$, respectively. Hence we have to choose the displacement d large enough to ensure that

$$A(F+d) > \text{PricePayerSwaption}(T, K_i)$$

and/or

$$A(K_i+d) > \text{PriceReceiverSwaption}(T, K_i)$$

is satisfied for all i.

Inserting the normal formulas (2.5) and (2.6), both equations boil down to the same condition

$$F + \Delta_i + d > \Delta_i \Phi\left(\frac{\Delta_i}{\sigma_{N,i}\sqrt{T}}\right) + \sigma_{N,i}\sqrt{T}\phi\left(\frac{\Delta_i}{\sigma_{N,i}\sqrt{T}}\right) \tag{3.10}$$

where $\sigma_{N,i}$ denotes normal volatility for the offset Δ_i.[12]

Consequently we need to choose the displacement d such that

$$d > \max_i \left\{\Delta_i \Phi\left(\frac{\Delta_i}{\sigma_{N,i}\sqrt{T}}\right) + \sigma_{N,i}\sqrt{T}\phi\left(\frac{\Delta_i}{\sigma_{N,i}\sqrt{T}}\right) - \Delta_i\right\} - F.$$

[10] The condition is obviously satisfied for $\tilde{d} \geq d$.
[11] Note that for SABR we only consider the case $\beta > 0$.
[12] Note that by setting $\Delta_i = 0$, one recovers the ATM formula (3.7) as a special case.

3.4 Excursion2: Transforming volatilities for highly correlated underlyings

Usually in a market only swaptions with a certain underlying float rate tenor are quoted. So how do we value and hedge swaptions with a different floating leg convention? A wide-spread application for this are EUR single callable swaps paying fix against 3 months floating Libor which deviates from the market standard of 6 months.[13] Another example are single callable bonds. They can be decomposed in a non callable bond and a swaption on an underlying exchanging a fixed rate against a floating rate plus a bond or funding spread. In both examples the associated forward swaprate $\tilde{F}$ deviates from the market standard forward swaprate F and the difference $b(t) = \tilde{F}(t) - F(t)$ is called basis spread. The question arises which volatility can be used to price the non-standard swaption on the underlying $\tilde{F}$.

A common assumption is to postulate a constant basis spread b.[14] Consider a swaption with maturity T, strike $\tilde{K}$ and forward swaprate $\tilde{F}$ and let $\tilde{\sigma}_N(\tilde{F}, \tilde{K})$ and $\tilde{\sigma}_{LN}(\tilde{F}, \tilde{K})$ be the unknown normal and lognormal implied volatilities, respectively. Dropping A and T to ease the notation one gets the following relation in the lognormal case:

$$\begin{aligned} C_{LN}(\tilde{F}, \tilde{K}, \tilde{\sigma}_{LN}(\tilde{F}, \tilde{K})) &= A \cdot \mathbb{E}^{Q_A}[\tilde{F}(T) - \tilde{K}]^+ = A \cdot \mathbb{E}^{Q_A}[(F(T) + b) - (K + b)]^+ \\ &= A \cdot \mathbb{E}^{Q_A}[F(T) - K]^+ = C_{LN}(F, K, \sigma_{LN}(F, K)), \end{aligned} \tag{3.11}$$

where $K := \tilde{K} - b$ and $\sigma_{LN}(F, K)$ is the quoted lognormal market volatility at the adjusted strike K. Here we implicitly assumed that both F and $\tilde{F}$ have the same annuity.[15] Analogously for the normal case:

$$C_N(\tilde{F}, \tilde{K}, \tilde{\sigma}_N(\tilde{F}, \tilde{K})) = C_N(F, K, \sigma_N(F, K))$$

where $\sigma_N(F, K)$ is the quoted normal market volatility at the adjusted strike K.

The normal formula (2.5) remains invariant if one shifts both forward rate and strike by the same rate, i.e.

$$C_N(F, K, \sigma_N(F, K)) = C_N(\tilde{F}, \tilde{K}, \sigma_N(F, K)).$$

Thus, the normal volatilities of F and $\tilde{F}$ are the same if one adjusts the strike according to the basis b:

$$\tilde{\sigma}_N(\tilde{F}, \tilde{K}) = \sigma_N(F, K) \tag{3.12}$$

Let $m(F, K) = F - K$ be the (absolute) moneyness. Since $m(F, K) = m(\tilde{F}, \tilde{K}) = m$ one can rewrite the last formula in terms of moneyness as

$$\tilde{\sigma}_N(m) = \sigma_N(m)$$

meaning that the non-standard volatility $\tilde{\sigma}_N$ is equal to the market volatility σ_N taken at the same level of (absolute) moneyness.

For the lognormal case the calculation is a bit more involved. Obviously one way is to numerically solve (3.11) for $\tilde{\sigma}_{LN}(\tilde{F}, \tilde{K})$. Alternatively one can use (3.12) equipped with a

[13] At least for maturities > 1 year.

[14] Assuming a constant spread is in fact the same as assuming a deterministic spread as $F(t)$ and $\tilde{F}(t)$ are both martingales:

$$F(0) + b(0) = \tilde{F}(0) = \mathbb{E}^{Q_A}[\tilde{F}(t)] = \mathbb{E}^{Q_A}[F(t) + b(t)] = F(0) + b(t)$$

[15] This holds for example for collateralized $3m$ and $6m$ EUR swaptions where the fix leg frequency is always yearly and the annuities are calculated on the same discounting curve. If $\tilde{F}$ is however an unsecured forward rate with the annuity being computed on a funding curve, then a change of measure from the unsecured to the secured annuity measure is required.

closed-form approximation formula to convert between lognormal and normal volatilities. For example using (3.2) one gets

$$\tilde{\sigma}_{LN}(\tilde{F},\tilde{K})\cdot\frac{\tilde{F}+\tilde{K}}{2}\approx\tilde{\sigma}_N(\tilde{F},\tilde{K})=\sigma_N(F,K)\approx\sigma_{LN}(F,K)\cdot\frac{F+K}{2}$$

and hence

$$\tilde{\sigma}_{LN}(\tilde{F},\tilde{K})\approx\sigma_{LN}(F,K)\frac{F+K}{\tilde{F}+\tilde{K}}. \tag{3.13}$$

For $F=K$ this reduces to

$$\tilde{\sigma}_{LN}(\tilde{F},\tilde{K})\approx\sigma_{LN}(F,K)\frac{F}{\tilde{F}} \tag{3.14}$$

We explicitly warn that formulas (3.13) and (3.14) are only valid for sufficiently small T and $\frac{K}{F}\approx 1$. In the ATM case the inaccuracy in the time dimension can be avoided by using the exact transformation (3.5) to obtain

$$\tilde{\sigma}_{LN}(\tilde{F},\tilde{F})=\frac{2\sqrt{2}}{\sqrt{T}}\cdot erf^{-1}\left(\frac{F}{\tilde{F}}\cdot erf\left(\frac{\sigma_{LN}(F,F)\sqrt{T}}{2\sqrt{2}}\right)\right) \tag{3.15}$$

One should however bear in mind that even this exact formula is based on the assumption of the basis spread being constant.

4 Relationship between (displaced) lognormal and normal Greeks

In this section we use the results of section 3 to derive explicit formulas linking the "new" normal and displaced lognormal sensitivities to the "old" lognormal ones. After elaborating on the set-up in section 4.1, section 4.2 sheds some light on the potential traps that need to be circumvented to ensure that the prices of interest rate products are indeed invariant to the particular choice of the volatility market data representation - as this prerequisite is required to derive the formulas to follow. Section 4.3 discusses several potential choices of delta and vega risk factors and their role when transforming sensitivities between the different volatility input regimes. Here we also explain what we mean by the terms lognormal, normal or displaced lognormal delta/vega. Section 4.4 derives linearisation of the implicit dependence between the volatilities. Based on this, section 4.5 states the analytical formulas to link deltas and vegas in a lognormal to those in a normal or displaced lognormal volatility market data representation.

4.1 Set-up

To price and model the risk of an interest rate derivative one needs to calibrate a volatility model, flat or equipped with smile, to the (displaced) lognormal or normal volatility cube quoted by the broker consisting of a set of option expiries $T_1,\ldots,T_n$ and swap tenors $\tilde{T}_1,\ldots,\tilde{T}_{\tilde{n}}$. We are, thus, dealing with a $n\times\tilde{n}$ matrix of vectors (a "cube") $\Sigma=(\sigma_{ij})$, where each component $\sigma_{ij}=[\sigma_{ij}(1),\ldots,\sigma_{ij}(m)]$ is a vector of m smile volatilities with corresponding relative strike offsets $M_{ij}=[M_{ij}(1),\ldots,M_{ij}(m)]$.[16,17] The corresponding absolute strikes for the i,j-th smile with corresponding forward swaprate F_{ij} are given by $K_{ij}=[K_{ij}(1),\ldots,K_{ij}(m)]$ where $K_{ij}(k)=F_{ij}+M_{ij}(k)$. $\sigma_{ij}(k)$, thus, corresponds to a market traded swaption with maturity T_i, tenor $\tilde{T}_j$ and moneyness $M_{ij}(k)$. Clearly this cube structure requires three

[16] For the sake of notational simplicity we implicitly assumed that we have the same number of tenors for every maturity and the same number of smile volatilities at each maturity/tenor bucket.
[17] It is common to use relative strikes in the form of moneyness and we stick to this convention.

indices. To ease the notation we enumerate the maturity/tenor buckets by a single index i taking values from 1 to $n \times \tilde{n}$. This way we have a sequence of market smiles $\sigma_i := [\sigma_i(1), \ldots, \sigma_i(m)]$ with moneyness strikes $M_i := [M_i(1), \ldots, M_i(m)]$ and corresponding forward swaprate F_i.

In what follows we will be using $\Sigma_N = (\sigma_{N,i})_{i=1,\ldots,n\times\tilde{n}}$, $\Sigma_{LN} = (\sigma_{LN,i})_{i=1,\ldots,n\times\tilde{n}}$ and $\Sigma_{LN}^d = (\sigma_{LN,i}^d)_{i=1,\ldots,n\times\tilde{n}}$ to denote the normal, lognormal and displaced lognormal volatility cube.

4.2 Invariance of prices when changing the volatility representation

All other factors being the same, the price of an interest rate product is invariant to the choice of the volatility market data representation. The smile model of choice, let it be explicit or numerical, has to be calibrated to the quoted volatility cube. Without loss of generality we assume that the broker collects option prices from which she then derives implied (displaced) lognormal and implied normal volatilities using (2.1) - (2.6). Suppose she publishes both normal as well as (displaced) lognormal data and forwards.

A naive but intuitive example of a smile model is to take the (displaced) lognormal volatilities and create a (displaced) lognormal surface by interpolating linearly in both the strike as well as the time dimension.[18] Obviously, one can easily "calibrate" this model by simply taking the quoted (displaced) lognormal volatilities. However, one may as well collect the normal volatilities instead and use them as the model input. In this case the system has to internally convert the normal input into the required (displaced) lognormal volatilities using formulas (2.1) - (2.6). As long as the conversion algorithm is exactly the same as the one applied by the broker, both ways will trivially result in identical volatility surfaces and, thus, in the same prices for all interest rate products.[19]

There are however trivial but important limitations to this view which one needs to bear in mind. As discussed in section 3.2, while it is always possible to transform a swaption price into a normal volatility, this does not hold in general for a (displaced) lognormal model. Especially in low rate market environments option prices are often found to exceed the no-arbitrage bounds implied by a (displaced) lognormal model and the corresponding buckets need to be dropped from the quoted volatility matrix. In this case it obviously matters whether one has collected the dense normal or the sparse (displaced) lognormal data set from the broker in the first place. Moreover, one has to be careful not to unintentionally change the smile model when changing the volatility input. This holds in particular for the SABR model or, to be more precise, for the set of different asymptotic formulas that are typically taken as analytic approximations of the SABR model. The lognormal and the normal formulas are subject to different approximation errors. Therefore, if one uses a lognormal formula for lognormal input and a normal formula for normal input one always implicitly changes the smile model along with the type of the input volatility. It is therefore advisable to stick to one type of approximation, preferably normal, and, if required, convert any input data beforehand. Lastly, thinking about the very simple smile model from above, it should be noted that of course it is not the same to interpolate volatilities and convert them thereafter or to convert the volatilities first and apply the interpolation afterwards.

In the following we will assume that the price of an interest rate product is fully invariant to the particular choice of the volatility market data representation.

4.3 The delta and vega risk factor coordinate systems

When deriving a sensitivity as a partial derivative one has to decide which variable to bump and which others to keep constant during the calculation. These "risk factors" span a coordi-

[18] We warn to use this approach in practice as it does not guarantee an arbitrage-free volatility surface.
[19] In particular the forward used and published by the broker should be used.

nate system which is supposed to cover all dimensions of uncertainty/risk that can be handled within the model.[20] In the followin we will however solely concentrate on the delta and vega dimension.

In section 4.5 we analyse two possible vega coordinate system specifications. We will start in section 4.5.1 with the theoretically interesting case of taking every individual cube volatility (of whatever type) as a vega risk factor.[21] This implicitly assumes that the definition of the risk factors always corresponds to the type of the input volatility - implying that the representation of the (vega and delta) risk always changes alongside with the volatility quoting convention. Note, however, that this vega coordinate system raises the concern of how stable the smile reacts to shifts of individual volatilities (while keeping all others the same). Usually, one does not inter- and extrapolate the smile directly but uses some parametrization like the SABR model. The transition from an individual volatility bump to a change in the parametrization can potentially be unstable and noisy.

A more natural choice is to take the parameters of the smile model/parametrization directly as the vega risk factors where we assume, as it is market standard, that the ATM volatility is always part of this system. An intuitive example of this is the SABR model with α being replaced by the ATM volatility. We follow this route in section 4.5.2 bu we will only concentrate on the ATM sensitivities while assuming that the other parts of the smile parametrization are independent of the type of the input volatility. Analogous to before we assume that the definition of the ATM risk factors is always in line with volatility quoting convention.[22]

In either coordinate system we call the deltas and vegas lognormal/displaced lognormal/normal if the vega risk factors are lognormal/displaced lognormal/normal (cube or ATM) volatilities.

When it comes to the delta, a natural choice is to take the quotes of the calibration products used to construct the entire curve universe as risk factors as they are typically also used as hedge trades. In this coordinate system the delta is calculated by bumping one market quote at a time while freezing all the others, implying that the whole bootstrapping procedure has to be repeated whenever a sensitivity is calculated.

A different possibility is to store each curve independently via a set of suitable building blocks - like zero-rates, discount factors or (forward) swaprates. Whenever one pillar of a curve is bumped all other building blocks of this and all other curves are assumed to remain unchanged. Choosing e.g. swaprates has the advantage that each forward delta risk factor only influences one smile - as long as the swaprates are in line with the swaptions and cap/floors used in the calibration. Note however that using swaprates as delta risk factors is a mere theoretical concept and should be treated with great care in practice. Taking all the swaprates corresponding to the entire volatility cube is, for example, an inadmissible choice as this would lead to collinearity in the risk factor system. A valid choice are non overlapping adjacent swaprates that span the whole delta risk dimension.

A related choice is to specify only one base interest rate curve per currency (with suitable building blocks) and define all other curves, within the same currency, as spreads over this curve. A natural choice is to take the standard-tenor LIBOR curve as the base curve. This has the convenient feature that only a small subset of the delta risk factors, namely the building

[20] There risks, like e.g. liquidity or close-out risk, that are acknowledged but cannot be captured by standard pricing models like Black-Scholes-Merton, SABR, Libor Market Model et cetera. As a result these risk dimensions cannot be part of the risk factor coordinate system of these models and must be dealt with externally.

[21] Basically in this set-up the smile model/parametrization (e.g. SABR) is recalibrated every time some curve input or one of the cube volatilities is bumped.

[22] We want to emphasize that this does not have to be the case. One could for example always use the displaced lognormal volatilities (corresponding to the displacement used in the SABR model) as the ATM risk factors - independent of how the volatilities are quoted on the market. In this case and assuming the conditions of section 4.2 are satisfied, we would see no impact (at least on the deltas) if we were to change the volatility input type.

blocks of the base curve, trigger a volatility re-bootstrap. In this example the delta risk factor universe is given by the building blocks of the base curve plus all spreads.

As it is not the main focus of this study we restrict ourself to one delta coordinate system in the following - namely we follow the first route and fix the quotes of the calibration products used to construct the entire curve universe. In the following we denote $\mathcal{S} = (S_j)_j$ as the delta risk factor coordinate system where S_j is the quote of the jth calibration product.

We still haven't fully specified our risk coordinate system. When bumping a delta risk factor both vega specifications discussed above postulate the ATM volatility (of whatever type) to remain the same. However, when a forward changes so does the definition of ATM. So we are still left with the choice of a sticky strike or a sticky delta/moneyness regime. In this paper we always assume sticky moneyness volatilities when calculating the delta.

4.4 Normal to lognormal linearisation

As we have discussed in section 3, for an arbitrary moneyness no closed-form transformation between (displaced) lognormal and normal implied volatilities exists. However, using the implicit function theorem we can easily obtain explicit linearisation of the implicit dependence between the volatilities. These linearisation are required to obtain the analytical transformation formulas for converting lognormal to normal sensitivities of section 4.5. In this section we fix a single expiry T, swap tenor and moneyness M to avoid all the tedious subindices. Note that in the following we are using formulas (2.1) - (2.6) in terms of the moneyness M instead of the strike K, where by definition $M := K - F$.[23] By doing so we are assuming that the moneyness M instead of the strike K remains constant when calculating the derivative with respect to F (i.e. sticky moneyness).

As the quoted normal and (displaced) lognormal volatilities are just alternative representation of the same market option prices C_{mkt} we have[24]

$$C_{mkt} = C_{LN}(A, F, M, T, \sigma_{LN}) = C_N(A, F, M, T, \sigma_N) = C^d_{LN}(A, F, M, T, \sigma^d_{LN}),$$

which can be rewritten as

$$G_1(A, F, M, T, \sigma_N, \sigma^d_{LN}) := C_N(A, F, M, T, \sigma_N) - C^d_{LN}(A, F, M, T, \sigma^d_{LN}) = 0$$

and

$$G_2(A, F, M, T, \sigma^d_{LN}, \sigma_{LN}) := C^d_{LN}(A, F, M, T, \sigma^d_{LN}) - C_{LN}(A, F, M, T, \sigma_{LN}) = 0.$$

These equations implicitly define the transformation functions

$$\sigma^d_{LN} = \sigma^d_{LN}(A, F, M, T, \sigma_N) \text{ and } \sigma_{LN} = \sigma_{LN}(A, F, M, T, \sigma^d_{LN}). \tag{4.1}$$

For further reference we add the derivatives of the Black-Scholes and Bachelier formulas with respect to the forward in the moneyness coordinate system:

$$\frac{\partial}{\partial F} C^d_{LN}(A, F, M, T, \sigma^d_{LN}) = A\left(\Phi\left(q^{LN,d}_{+}\right) - \Phi\left(q^{LN,d}_{-}\right)\right)$$

$$\frac{\partial}{\partial F} C_N(F, M, T, \sigma_N) = 0 \tag{4.2}$$

$$\text{where } q^{LN,d}_{+/-} = \frac{\log(F+d) - \log(F+M+d)}{\sigma^d_{LN}\sqrt{T}} +/_- \frac{\sigma^d_{LN}\sqrt{T}}{2}$$

[23] For convenience we stick, however, to the same function names C_{LN}, C^d_{LN} and C_N.

[24] Here we implicitly assume that the market option prices are small enough to be representable by a lognormal model.

Even though we do not have the explicit form of the functions (4.1), using the Implicit Function Theorem we can calculate their first derivatives analytically. To ease the notation we drop the dependency on A, T and M in the following. But it is important always to remember that we are working in a moneyness and not a strike coordinate system.

$$
\begin{aligned}
\frac{\partial \sigma_{LN}^{d}}{\partial \sigma_N}(F,\sigma_N) &= \frac{\frac{\partial C_N}{\partial \sigma_N}}{\frac{\partial C_{LN}^{d}}{\partial \sigma_{LN}^{d}}}(F,\sigma_N) = \frac{\text{normal vega}}{\text{displ lognorm vega}} \\
&= \frac{\phi\left(\frac{M}{\sigma_N\sqrt{T}}\right)}{(F+d)\cdot\phi\left(q_{+}^{LN,d}\right)} \\
\\
\frac{\partial \sigma_{LN}}{\partial \sigma_{LN}^{d}}(F,\sigma_{LN}^{d}) &= \frac{\frac{\partial C_{LN}^{d}}{\partial \sigma_{LN}^{d}}}{\frac{\partial C_{LN}}{\partial \sigma_{LN}}}(F,\sigma_{LN}^{d}) = \frac{\text{displ lognorm vega}}{\text{lognorm vega}} \\
&= \frac{(F+d)\cdot\phi\left(q_{+}^{LN,d}\right)}{F\cdot\phi\left(q_{+}^{LN,0}\right)} \qquad (4.3) \\
\\
\frac{\partial \sigma_{LN}^{d}}{\partial F}(F,\sigma_N) &= \frac{\frac{\partial C_N}{\partial F}-\frac{\partial C_{LN}^{d}}{\partial F}}{\frac{\partial C_{LN}^{d}}{\partial \sigma_{LN}^{d}}} = \frac{\text{normal moneyness delta} - \text{displ lognorm moneyness delta}}{\text{displ lognorm vega}} \\
&= \frac{\Phi\left(q_{-}^{LN,d}\right)-\Phi\left(q_{+}^{LN,d}\right)}{(F+d)\cdot\sqrt{T}\cdot\phi\left(q_{+}^{LN,d}\right)} \\
\\
\frac{\partial \sigma_{LN}}{\partial F}(F,\sigma_{LN}^{d}) &= \frac{\frac{\partial C_{LN}^{d}}{\partial F}-\frac{\partial C_{LN}}{\partial F}}{\frac{\partial C_{LN}}{\partial \sigma_{LN}}} = \frac{\text{displ lognorm moneyness delta} - \text{lognorm moneyness delta}}{\text{lognorm vega}} \\
&= \frac{\Phi\left(q_{+}^{LN,d}\right)-\Phi\left(q_{-}^{LN,d}\right)-\Phi\left(q_{+}^{LN,0}\right)+\Phi\left(q_{-}^{LN,0}\right)}{F\cdot\sqrt{T}\cdot\phi\left(q_{+}^{LN,0}\right)}.
\end{aligned}
$$

Note that for $M \to 0$ (i.e. ATM) the formulas in (4.3) simplify to

$$
\begin{aligned}
\frac{\partial \sigma_{LN}^d}{\partial \sigma_N}(F,\sigma_N) &= \frac{1}{F+d} \cdot e^{\frac{1}{8}(\sigma_{LN}^d)^2 T} \\
\frac{\partial \sigma_{LN}}{\partial \sigma_{LN}^d}(F,\sigma_{LN}^d) &= \frac{F+d}{F} \cdot e^{\frac{1}{8}\left((\sigma_{LN})^2-(\sigma_{LN}^d)^2\right)T} \\
\frac{\partial \sigma_{LN}^d}{\partial F}(F,\sigma_N) &= \frac{1-2\cdot\Phi(\frac{1}{2}\sigma_{LN}^d\sqrt{T})}{(F+d)\cdot\sqrt{T}\cdot\phi\left(\frac{1}{2}\sigma_{LN}^d\sqrt{T}\right)} \\
&= -\sqrt{\frac{2\pi}{T}}\cdot\frac{1}{F+d}\cdot e^{\frac{1}{8}(\sigma_{LN}^d)^2 T}\cdot \operatorname{erf}\left(\frac{\sigma_{LN}^d\sqrt{T}}{2\sqrt{2}}\right) \\
\frac{\partial \sigma_{LN}}{\partial F}(F,\sigma_{LN}^d) &= \frac{2\cdot\Phi\left(\frac{1}{2}\sigma_{LN}^d\sqrt{T}\right)-2\cdot\Phi\left(\frac{1}{2}\sigma_{LN}\sqrt{T}\right)}{F\cdot\sqrt{T}\cdot\phi\left(\frac{1}{2}\sigma_{LN}\sqrt{T}\right)} \\
&= \sqrt{\frac{2\pi}{T}}\cdot\frac{1}{F}\cdot e^{\frac{1}{8}(\sigma_{LN})^2 T}\cdot\left[\operatorname{erf}\left(\frac{\sigma_{LN}^d\sqrt{T}}{2\sqrt{2}}\right)-\operatorname{erf}\left(\frac{\sigma_{LN}\sqrt{T}}{2\sqrt{2}}\right)\right]
\end{aligned}
\tag{4.4}
$$

Obviously the equations (4.4) can also be derived directly from Proposition 3.2.

4.5 Delta and vega transformations

Using the formulas from section 4.4 we now have a look at what influence the volatility quoting convention has on the delta and vega sensitivities. We do this in the context of the two vega risk factor coordinate systems described in section 4.3 while always taking the calibration product quotes of the forward curve as the delta risk factors.

In section 4.5.1 we take all (ATM as well as offset) volatilities of the cube as the vega risk factors. Here we analyse how the vega changes when it is based on an elementwise bumping of the normal or displaced lognormal cube ("displaced lognormal/normal vega") instead of the lognormal volatility cube ("lognormal vega"). Similarly we investigate the effect on the delta when it is calculated under the premise of keeping all entries of the normal or displaced lognormal cube ("displaced lognormal/normal delta") instead of the lognormal cube ("lognormal delta") constant.

In section 4.5.2 we follow the rather generic approach by taking the ATM volatilities plus some not further specified variables parametrizing the smile as the vega risk factors.[25] Here we investigate how the deltas and vegas change when taking the normal or displaced lognormal ATM volatilities ("displaced lognormal/normal delta and vega") instead of the lognormal ones ("lognormal delta and vega") as the ATM vega risk factors. The rest of the smile model/parametrization is assumed to remain fixed when the volatility quoting convention is changed.[26] Note that the system of section 4.5.1 that takes all entries of the volatility cube as

[25] Think about the SABR model with α being replaced by the ATM volatility.

[26] We want to explain this crucial assumption of "keeping the smile model/parametrization unchanged" in more detail using a few examples:
(1) For a SABR smile parametrization the assumption implies for example that we are always using the same SABR approximation formula and, thus, have the same calibrated parameters α, β, ρ and ν. What is now different when the normal or displaced lognormal ATM volatilities are fixed instead of the lognormal ones? We know that $\alpha = \alpha(F,\sigma_{LN}) \approx \frac{\sigma_{LN}}{F^{\beta-1}}$, $\alpha = \alpha(F,\sigma_N) \approx \frac{\sigma_N}{F^\beta}$ and $\alpha = \alpha(F,\sigma_{LN}^d) \approx \sigma_{LN}^d \frac{F+d}{F^\beta}$. Obviously the parameter α, reacts differently to a 1bp change in σ_{LN} than it does to a 1bp change in σ_N or σ_{LN}^d. Similar the influence of F on α varies depending on which type of ATM volatility is kept fixed.
(2) Another admissible choice would be to take ATM volatilities (lognormal, normal or displaced lognormal) plus volatility offsets. Here the smile model/parametrization is only left unchanged if the offset definition remains the same independent of the ATM volatility type (i.e. one must not take lognormal ATM volatilities plus lognormal offsets vs. normal ATM volatilities plus normal offsets).

the vega risk factors violates this assumption.

In the following we will make use of the fact that any forward swaprate F can be expressed in terms of a long and a short spot swaprate, S_l and S_s, as

$$F = \frac{A_l}{A} S_l - \frac{A_s}{A} S_s, \tag{4.5}$$

where A is the forward annuity corresponding to the forward swaprate F, and A_l and A_s are the annuities of the long and the short spot swaprates S_l and S_s, respectively. If the long swaprate S_l and short swaprate S_s happen to be liquid quotes in the market and part of the delta risk factor coordinate system $\mathcal{S}$, then one has:

$$\sum_{\{j: S_j \in \mathcal{S}\}} \frac{\partial F}{\partial S_j} = \frac{A_l}{A} - \frac{A_s}{A} = 1. \tag{4.6}$$

Note however that this is not the case in general. For example a 7Y+10Y forward rate can be decomposed into a long 17 years and short 7 years spot swaprate. However, there is no liquid quote for a swap maturing in 17 years and the corresponding swaprate will not be a delta risk factor. But depending on the curve construction and interpolation method, (4.6) will still hold approximately:

$$\sum_{\{j: S_j \in \mathcal{S}\}} \frac{\partial F}{\partial S_j} \approx 1. \tag{4.7}$$

4.5.1 The volatility cube case

In this subsection we consider a normal pricing function $V_N(\mathcal{S}, \Sigma_N)$ and a (displaced) lognormal pricing function $V_{LN}^d(\mathcal{S}, \Sigma_{LN}^d)$, where $\mathcal{S} = (S_j)_j$ is the vector of the forward curve calibration product quotes and Σ_N and Σ_{LN}^d denote the normal and (displaced) lognormal volatility cubes. As discussed in section 4.2, we assume that the model and its calibration is independent of the particular choice of the volatility market data representation, that is

$$V_N(\mathcal{S}, \Sigma_N) = V_{LN}(\mathcal{S}, \Sigma_{LN})$$

and

$$V_{LN}^d(\mathcal{S}, \Sigma_{LN}^d) = V_{LN}(\mathcal{S}, \Sigma_{LN})$$

The change between normal and lognormal first order sensitivities amounts to calculating the gradients of the pricing function in different coordinate systems, i.e. $(\mathcal{S}, \Sigma_N)$ vs. $(\mathcal{S}, \Sigma_{LN})$ or $(\mathcal{S}, \Sigma_{LN}^d)$ vs. $(\mathcal{S}, \Sigma_{LN})$.

Delta transformation: In the risk factor system $(\mathcal{S}, \Sigma_N)$ the delta is calculated by bumping calibration product quotes S while keeping all volatilities of the normal cube Σ_N fixed. With the chain rule we get:

$$\begin{aligned}\underbrace{\frac{\partial V_N}{\partial S_j}(\mathcal{S}, \Sigma_N)}_{\text{normal delta}} = \frac{\partial V_{LN}}{\partial S_j}(\mathcal{S}, \Sigma_{LN}(F(\mathcal{S}), \Sigma_N)) = \underbrace{\frac{\partial V_{LN}}{\partial S_j}(\mathcal{S}, \Sigma_{LN})}_{\text{lognormal delta}} + \\ \sum_{i=1}^{n\times\tilde{n}} \sum_{k=1}^{m} \underbrace{\frac{\partial V_{LN}}{\partial \sigma_{LN,i}(k)}(\mathcal{S}, \Sigma_{LN})}_{\text{lognormal vega}} \times \frac{\partial \sigma_{LN,i}(k)}{\partial F_i}(F_i(\mathcal{S}), \sigma_{N,i}(k)) \times \frac{\partial F_i}{\partial S_j}(\mathcal{S})\end{aligned} \tag{4.8}$$

and

$$\underbrace{\frac{\partial V_{LN}^d}{\partial S_j}(\mathcal{S},\Sigma_{LN}^d)}_{\text{displaced lognormal delta}} = \frac{\partial V_{LN}}{\partial S_j}(\mathcal{S},\Sigma_{LN}(F(\mathcal{S}),\Sigma_{LN}^d)) = \underbrace{\frac{\partial V_{LN}}{\partial S_j}(\mathcal{S},\Sigma_{LN})}_{\text{lognormal delta}} + \sum_{i=1}^{n\times\tilde{n}}\sum_{k=1}^{m}\underbrace{\frac{\partial V_{LN}}{\partial \sigma_{LN,i}(k)}(\mathcal{S},\Sigma_{LN})}_{\text{lognormal vega}} \times \frac{\partial \sigma_{LN,i}(k)}{\partial F_i}(F_i(\mathcal{S}),\sigma_{LN,i}^d(k)) \times \frac{\partial F_i}{\partial S_j}(\mathcal{S}), \tag{4.9}$$

where we have used the fact that $\frac{\partial \sigma_{LN,i}^d(k)}{\partial F_p} = 0$ for $p \neq i$. Note that the term $\frac{\partial F_i}{\partial S_j}(\mathcal{S})$ solely depends on the curve bootstrapping algorithm and is independent from the volatilities. Due to the recursive nature of the bootstrap and the ceteris-paribus definition of a sensitivity it will be zero for many i.[27]

Equations (4.8) and (4.9) show that the displaced lognormal or normal delta is equal to the lognormal one adjusted by a term linked to the lognormal vega. Zero-vega products, thus, always have the same delta, independent of the volatility input type. This however also implies that depending on the volatility market data representation usually a different number of delta hedge trades (with zero vega) is required to hedge the delta risk of a volatility sensitive product, i.e. *the delta hedge ratios will in general change*.

Vega transformation: Analogously we obtain the link between normal and lognormal or between displaced lognormal and lognormal vegas:

$$\begin{aligned}\underbrace{\frac{\partial V_N}{\partial \sigma_{N,i}(k)}(\mathcal{S},\Sigma_N)}_{\text{normal vega}} &= \frac{\partial V_{LN}}{\partial \sigma_{N,i}(k)}(\mathcal{S},\Sigma_{LN}(\mathcal{S},\Sigma_N)) \\ &= \underbrace{\frac{\partial V_{LN}}{\partial \sigma_{LN,i}(k)}(\mathcal{S},\Sigma_{LN})}_{\text{lognormal vega}} \times \frac{\partial \sigma_{LN,i}(k)}{\partial \sigma_{N,i}(k)}(F_i(\mathcal{S}),\sigma_{N,i}(k)),\end{aligned} \tag{4.10}$$

$$\begin{aligned}\underbrace{\frac{\partial V_{LN}^d}{\partial \sigma_{LN,i}^d(k)}(\mathcal{S},\Sigma_{LN}^d)}_{\text{displaced lognormal vega}} &= \frac{\partial V_{LN}}{\partial \sigma_{LN,i}^d(k)}(\mathcal{S},\Sigma_{LN}(\mathcal{S},\Sigma_{LN}^d)) \\ &= \underbrace{\frac{\partial V_{LN}}{\partial \sigma_{LN,i}(k)}(\mathcal{S},\Sigma_{LN})}_{\text{lognormal vega}} \times \frac{\partial \sigma_{LN,i}(k)}{\partial \sigma_{LN,i}^d(k)}(F_i(\mathcal{S}),\sigma_{LN,i}^d(k)),\end{aligned} \tag{4.11}$$

where we recognized that the kth lognormal volatility of the ith smile $\sigma_{LN,i}(k)$ depends on the kth normal/displaced lognormal volatility of the ith smile $\sigma_{N,i}(k)/\sigma_{LN,i}^d(k)$ but none of the other volatilities.

Equations (4.10) and (4.11) show that the displaced lognormal or normal vega is equal to the scaled lognormal vega. The influence on products and their vega hedge trades is exactly the same and, thus, *the vega hedge ratios are invariant to a change in the volatility market data*

[27] As an example assume $\mathcal{S} = (..., S_{5Y}, S_{7Y}, S_{10Y}, ...)$. Then the derivative $\frac{\partial F_i}{\partial S_{7Y}}(S)$ will only be non-zero for those forward swaprates that start or mature in $]5Y, 10Y[$, e.g. 1Y+5Y or 7Y+10Y.

representation.

Plugging in the expressions for the results from (4.3) into (4.8) - (4.11) we get explicit relations linking the deltas and vegas of the normal volatility cube risk factor system $(\mathcal{S}, \Sigma_N)$ with those of the (displaced) lognormal volatility cube risk factor system $(\mathcal{S}, \Sigma^d_{LN})$.

4.5.2 The at-the-money volatility case

In this subsection we consider a normal pricing function $V_N(\mathcal{S}, \Sigma_N|_{ATM})$ and a (displaced) lognormal pricing function $V^d_{LN}(\mathcal{S}, \Sigma^d_{LN}|_{ATM})$, where $\mathcal{S} = (S_j)_j$ is again the vector of calibration product quotes used to bootstrap the forward curve, $\Sigma_N|_{ATM}$ and $\Sigma^d_{LN}|_{ATM}$ denote the restrictions of the normal and the (displaced) lognormal cube to the ATM volatility matrix. Again we assume that the model and its calibration is independent of the particular choice of the volatility market data representation, that is

$$V_N(\mathcal{S}, \Sigma_N|_{ATM}) = V_{LN}(\mathcal{S}, \Sigma_{LN}|_{ATM})$$

and

$$V^d_{LN}(\mathcal{S}, \Sigma^d_{LN}|_{ATM}) = V_{LN}(\mathcal{S}, \Sigma_{LN}|_{ATM})$$

Delta transformation: In the risk factor systems $(\mathcal{S}, \Sigma_N|_{ATM})$ the delta is calculated by bumping calibration product quotes S while keeping all volatilities of the normal ATM matrix $\Sigma_N|_{ATM}$ and all smile parameters fixed. Using the results from (4.4), we get:

$$\begin{aligned}
\underbrace{\frac{\partial V_N}{\partial S_j}(\mathcal{S}, \Sigma_N|_{ATM})}_{\text{normal ATM delta}} &= \frac{\partial V_{LN}}{\partial S_j}(\mathcal{S}, \Sigma_{LN}|_{ATM}(F(\mathcal{S}), \Sigma_N|_{ATM})) \\
&= \underbrace{\frac{\partial V_{LN}}{\partial S_j}(\mathcal{S}, \Sigma_{LN}|_{ATM})}_{\text{lognormal ATM delta}} + \sum_{i=1}^{n\times\tilde{n}} \underbrace{\frac{\partial V_{LN}}{\partial \sigma_{LN,i}}(\mathcal{S}, \Sigma_{LN}|_{ATM})}_{\text{lognormal ATM vega}} \\
&\times \underbrace{\frac{\partial \sigma_{LN,i}}{\partial F_i}(F_i(\mathcal{S}), \sigma_{N,i})}_{=-\sqrt{\frac{2\pi}{T_i}}\cdot\frac{1}{F_i}\cdot e^{\frac{1}{8}(\sigma_{LN,i})^2 T_i}\cdot \mathrm{erf}\left(\frac{\sigma_{LN,i}\sqrt{T_i}}{2\sqrt{2}}\right)} \times \frac{\partial F_i}{\partial S_j}(\mathcal{S})
\end{aligned} \tag{4.12}$$

The only unknowns are the terms $\frac{\partial F_i}{\partial S_j}$ for all i and j. They can be easily computed with the curve construction library. Alternatively one can sum over all bucket deltas and use formula (4.7) to approximately cancel them out. Thereby one obtains a simple formula for the parallel delta:

$$\begin{aligned}
\underbrace{\sum_{\{j:S_j\in\mathcal{S}\}} \frac{\partial V_N}{\partial S_j}}_{\text{normal ATM parallel delta}} \approx \underbrace{\sum_{\{j:S_j\in\mathcal{S}\}} \frac{\partial V_{LN}}{\partial S_j}}_{\text{lognormal ATM parallel delta}} - \sum_{i=1}^{n\times\tilde{n}} \underbrace{\frac{\partial V_{LN}}{\partial \sigma_{LN,i}}}_{\text{lognormal ATM vega}} \\
\times \sqrt{\frac{2\pi}{T_i}} \cdot \frac{1}{F_i} \cdot e^{\frac{1}{8}(\sigma_{LN,i})^2 T_i} \cdot \mathrm{erf}\left(\frac{\sigma_{LN,i}\sqrt{T_i}}{2\sqrt{2}}\right)
\end{aligned} \tag{4.13}$$

In a similar fashion we obtain the link between displaced lognormal ATM parallel delta and lognormal ATM parallel delta:

$$\underbrace{\sum_{\{j:S_j\in\mathcal{S}\}}\frac{\partial V^d_{LN}}{\partial S_j}}_{\text{displ lognormal ATM parallel delta}} \approx \underbrace{\sum_{\{j:S_j\in\mathcal{S}\}}\frac{\partial V_{LN}}{\partial S_j}}_{\text{lognormal ATM parallel delta}} + \sum_{i=1}^{n\times\tilde{n}}\underbrace{\frac{\partial V_{LN}}{\partial\sigma_{LN,i}}}_{\text{lognormal ATM vega}}$$
$$\times\sqrt{\frac{2\pi}{T_i}}\cdot\frac{1}{F_i}\cdot e^{\frac{1}{8}(\sigma_{LN,i})^2T_i}\cdot\left[\operatorname{erf}\left(\frac{\sigma^d_{LN,i}\sqrt{T_i}}{2\sqrt{2}}\right)-\operatorname{erf}\left(\frac{\sigma_{LN,i}\sqrt{T_i}}{2\sqrt{2}}\right)\right] \tag{4.14}$$

Equations (4.13) and (4.14) show that the displaced lognormal or normal delta is equal to the lognormal one adjusted by a term linked to the lognormal vega. As before this shows that *the delta hedge ratios will in general change with the volatility market data representation*.

Vega transformation: The ATM vega relationships between lognormal and normal or lognormal and displaced lognormal are given by

$$\begin{aligned}\underbrace{\frac{\partial V_N}{\partial\sigma_{N,i}}(\mathcal{S},\Sigma_N|_{ATM})}_{\text{normal ATM vega}} &= \frac{\partial V_{LN}}{\partial\sigma_{N,i}}(\mathcal{S},\Sigma_{LN}|_{ATM}(F(\mathcal{S}),\Sigma_N|_{ATM}))\\ &= \underbrace{\frac{\partial V_{LN}}{\partial\sigma_{LN,i}}(\mathcal{S},\Sigma_{LN}|_{ATM})}_{\text{lognormal ATM vega}}\times\underbrace{\frac{\partial\sigma_{LN,i}}{\partial\sigma_{N,i}}(F_i(\mathcal{S}),\sigma_{N,i})}_{=\frac{1}{F_i}\cdot e^{\frac{1}{8}(\sigma_{LN,i})^2T_i}}\end{aligned} \tag{4.15}$$

and

$$\begin{aligned}\underbrace{\frac{\partial V^d_{LN}}{\partial\sigma^d_{LN,i}}(\mathcal{S},\Sigma^d_{LN}|_{ATM})}_{\text{displaced lognormal ATM vega}} &= \frac{\partial V_{LN}}{\partial\sigma^d_{LN,i}}(\mathcal{S},\Sigma_{LN}|_{ATM}(F(\mathcal{S}),\Sigma^d_{LN}|_{ATM}))\\ &= \underbrace{\frac{\partial V_{LN}}{\partial\sigma_{LN,i}}(\mathcal{S},\Sigma_{LN}|_{ATM})}_{\text{lognormal ATM vega}}\times\underbrace{\frac{\partial\sigma_{LN,i}}{\partial\sigma^d_{LN,i}}(F_i(\mathcal{S}),\sigma^d_{LN,i})}_{=\frac{F_i+d_i}{F_i}\cdot e^{\frac{1}{8}\left((\sigma_{LN,i})^2-(\sigma^d_{LN,i})^2\right)T_i}}\end{aligned} \tag{4.16}$$

Equations (4.15) and (4.16) show that the displaced lognormal or normal ATM vega is equal to the scaled lognormal ATM vega. Similar to before, *the ATM vega hedge ratios are invariant to a change in the volatility market data representation*.

5 Conclusion

In this paper we present formulas to link volatilities as well as vegas and deltas of a lognormal to those of a normal or displaced lognormal volatility market data representation. We provide a summary and a derivation of the most important transformation formulas between lognormal, displaced lognormal and normal volatilities. Furthermore we have shown that most transformations, like the well-known small maturity ATM formula (3.3), are only approximations and their implications can be poor and dangerous to rely on.

To circumvent these problems we based our greek transformations either on exact ATM formulas or on analytical linearisation between the volatilities. Especially the vega coordinate system described in section 4.5.2 is very generic and should cover the majority of cases in practice. The derived results can be used as a consistency check when switching the volatility input and as means to adapt the delta and vega risk limits in a sensible way. Although their derivation relies only on rather simple maths, they seem to be new (at least for the non-ATM case described in section 4.5.1) and not well known.

References

[1] Cyril Grunspan. Asymptotic expansions of the lognormal implied volatility: A model free approach. *arXiv, 8 Dec 2011*, 2011.

[2] Cyril Grunspan. A note on the equivalence between the normal and the lognormal implied volatility: A model free approach. *arXiv, 9 Dec 2011*, 2011.

[3] Peter Jaeckel. By implication. *http://www.pjaeckel.webspace.virginmedia.com/ByImplication.pdf*, 2010.

[4] J. Teichmann W. Schachermayer. How close are the option pricing formulas of bachelier and black-merton-scholes? *Mathematical Finance*, 18(1):155–170, 2008.