# Managing Smile Risk

**Article** *in* Wilmott · January 2002



**4 authors**, including:

Patrick S Hagan
XBTO Inc.


Andrew Lesniewski
City University of New York - Bernard M. Baruch College

# MANAGING SMILE RISK

PATRICK S. HAGAN*, DEEP KUMAR†, ANDREW S. LESNIEWSKI‡, AND DIANA E. WOODWARD§

**Abstract.** Market smiles and skews are usually managed by using local volatility models *a la* Dupire. We discover that the dynamics of the market smile predicted by local vol models is opposite of observed market behavior: when the price of the underlying decreases, local vol models predict that the smile shifts to higher prices; when the price increases, these models predict that the smile shifts to lower prices. Due to this contradiction between model and market, delta and vega hedges derived from the model can be unstable and may perform worse than naive Black-Scholes' hedges.

To eliminate this problem, we derive the *SABR* model, a stochastic volatility model in which the forward value satisfies

$$d\hat{F} = \hat{a}\hat{F}^{\beta} dW_1$$
$$d\hat{a} = \nu \hat{a}\, dW_2$$

and the forward $\hat{F}$ and volatility $\hat{a}$ are correlated: $dW_1 dW_2 = \rho dt$. We use singular perturbation techniques to obtain the prices of European options under the SABR model, and from these prices we obtain explicit, closed-form algebraic formulas for the implied volatility as functions of today's forward price $f = \hat{F}(0)$ and the strike $K$. These formulas immediately yield the market price, the market risks, including *vanna* and *volga* risks, and show that the SABR model captures the correct dynamics of the smile. We apply the SABR model to USD interest rate options, and find good agreement between the theoretical and observed smiles.

**Key words.** smiles, skew, dynamic hedging, stochastic vols, volga, vanna

**1. Introduction.** European options are often priced and hedged using Black's model, or, equivalently, the Black-Scholes model. In Black's model there is a one-to-one relation between the price of a European option and the volatility parameter $\sigma_B$. Consequently, option prices are often quoted by stating the *implied volatility* $\sigma_B$, the unique value of the volatility which yields the option's dollar price when used in Black's model. In theory, the volatility $\sigma_B$ in Black's model is a constant. In practice, options with different strikes $K$ require different volatilities $\sigma_B$ to match their market prices. See figure 1. Handling these market *skews* and *smiles* correctly is critical to fixed income and foreign exchange desks, since these desks usually have large exposures across a wide range of strikes. Yet the inherent contradiction of using different volatilities for different options makes it difficult to successfully manage these risks using Black's model.

The development of *local volatility models* by Dupire [2], [3] and Derman-Kani [4], [5] was a major advance in handling smiles and skews. Local volatility models are self-consistent, arbitrage-free, and can be calibrated to precisely match observed market smiles and skews. Currently these models are the most popular way of managing smile and skew risk. However, as we shall discover in section 2, the *dynamic* behavior of smiles and skews predicted by local vol models is exactly *opposite* the behavior observed in the marketplace: when the price of the underlying asset *decreases*, local vol models predict that the smile shifts to *higher* prices; when the price *increases*, these models predict that the smile shifts to *lower* prices. In reality, asset prices and market smiles move in the *same* direction. This contradiction between the model and the marketplace tends to de-stabilize the delta and vega hedges derived from local volatility models, and often these hedges perform worse than the naive Black-Scholes' hedges.

To resolve this problem, we derive the *SABR* model, a stochastic volatility model in which the asset price and volatility are correlated. Singular perturbation techniques are used to obtain the prices of European options under the SABR model, and from these prices we obtain a closed-form algebraic formula for the implied volatility as a function of today's forward price $f$ and the strike $K$. This closed-form formula for the implied volatility allows the market price and the market risks, including *vanna* and *volga* risks, to be

*phagan@bear.com; Bear-Stearns Inc., 383 Madison Avenue, New York, NY 10179
†BNP Paribas; 787 Seventh Avenue; New York NY 10019
‡BNP Paribas; 787 Seventh Avenue; New York NY 10019
§Societe Generale; 1221 Avenue of the Americas; New York NY 10020

obtained immediately from Black's formula. It also provides good, and sometimes spectacular, fits to the implied volatility curves observed in the marketplace. See figure 1.1. More importantly, the formula shows that the SABR model captures the correct dynamics of the smile, and thus yields stable hedges.

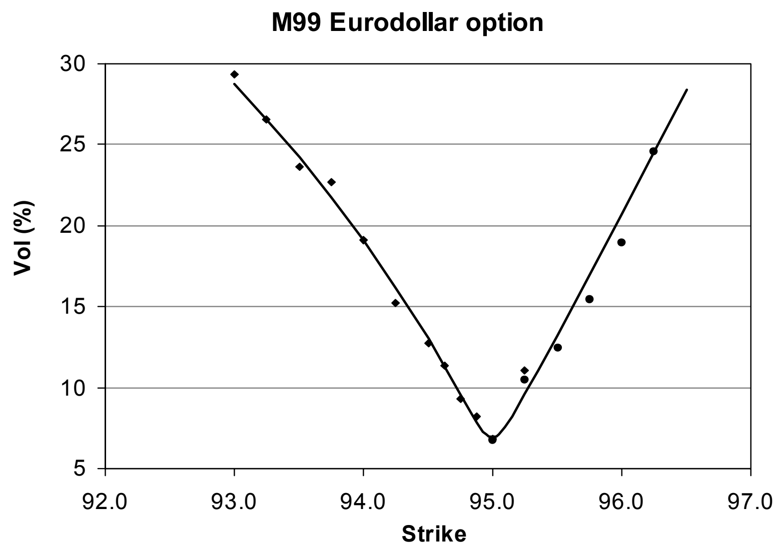


FIG. 1.1. *Implied volatility for the June 99 Eurodollar options. Shown are close-of-day values along with the volatilities predicted by the SABR model. Data taken from Bloomberg information services on March 23, 1999.*

**2. Reprise.** Consider a European call option on an asset $\mathcal{A}$ with *exercise date* $t_{ex}$, *settlement date* $t_{set}$, and *strike* $K$. If the holder exercises the option on $t_{ex}$, then on the settlement date $t_{set}$ he receives the underlying asset $\mathcal{A}$ and pays the strike $K$. To derive the value of the option, define $\hat{F}(t)$ to be the forward price of the asset for a forward contract that matures on the settlement date $t_{set}$, and define $f = \hat{F}(0)$ to be today's forward price. Also let $D(t)$ be the *discount factor* for date $t$; that is, let $D(t)$ be the value today of \$1 to be delivered on date $t$. In Appendix A the fundamental theorem of arbitrage free pricing [6], [8] is used to develop the theoretical framework for European options. There it is shown that the value of the call option is

$$V_{call} = D(t_{set}) \left\{ E[\hat{F}(t_{ex}) - K]^{+} | \mathfrak{F}_0 \right\}, \tag{2.1a}$$

and the value of the corresponding European put is

$$\begin{aligned} V_{put} &= D(t_{set}) E \left\{ [K - \hat{F}(t_{ex})]^{+} | \mathfrak{F}_0 \right\} \\ &\equiv V_{call} + D(t_{set})[K - f]. \end{aligned} \tag{2.1b}$$

Here the expectation $E$ is over the forward measure, and "$|\mathfrak{F}_0$" can be interpretted as "given all information available at $t = 0$." See Appendix A. In Appendix A it is also shown that the forward price $\hat{F}(t)$ is a Martingale under the forward measure. Therefore, the Martingale representation theorem implies that $\hat{F}(t)$ evolves according to

$$d\hat{F} = C(t, *) dW, \qquad \hat{F}(0) = f, \tag{2.1c}$$

for some coefficient $C(t,*)$, where $dW$ is Brownian motion in this measure. The coefficient $C(t,*)$ may be deterministic or random, and may depend on any information that can be resolved by time $t$. This is as far as the fundamental theory of arbitrage free pricing goes. In particular, one cannot determine the coefficient $C(t,*)$ on purely theoretical grounds. Instead one must postulate a mathematical *model* for $C(t,*)$.

European swaptions fit within an indentical framework. Consider a European swaption with exercise date $t_{ex}$ and fixed rate (strike) $R_{fix}$. Let $R_s(t)$ be the swaption's forward swap rate as seen at date $t$, and let $R_0 = \hat{R}_s(0)$ be the forward swap rate as seen today. In Appendix A we show that the value of a payer swaption is

$$V_{pay} = L_0 E\left\{[\hat{R}_s(t_{ex}) - R_{fix}]^+ | \mathfrak{F}_0\right\}, \tag{2.2a}$$

and the value of a receiver swaption is

$$\begin{aligned} V_{rec} &= L_0 E\left\{[R_{fix} - \hat{R}_s(t_{ex})]^+ | \mathfrak{F}_0\right\} \\ &\equiv V_{pay} + L_0[R_{fix} - R_0]. \end{aligned} \tag{2.2b}$$

Here $L_0$ is today's value of the *level* (*annuity*), which is a known quantity, and $E$ is the expectation over the *level measure* of Jamshidean [10]. In Appendix A it is also shown that the forward swap rate $\hat{R}_s(t)$ is a Martingale in this measure, so once again

$$d\hat{R}_s = C(t,*)dW, \qquad \hat{R}_s(0) = R_0, \tag{2.2c}$$

where $dW$ is Brownian motion. As before, the coefficient $C(t,*)$ may be deterministic or random, and cannot be determined from fundamental theory. Apart from notation, this is identical to the framework provided by equations 2.1a - 2.1c for European calls and puts. Caplets and floorlets can also be included in this picture, since they are just one period payer and receiver swaptions. For the remainder of the paper, we adopt the notation of 2.1a - 2.1c for general European options.

**2.1. Black's model and implied volatilities.** To go any further requires postulating a model for the coefficient $C(t,*)$. In [11], Black postulated that the coefficient $C(t,*)$ is $\sigma_B \hat{F}(t)$, where the *volatilty* $\sigma_B$ is a constant. The forward price $\hat{F}(t)$ is then geometric Brownian motion:

$$d\hat{F} = \sigma_B \hat{F}(t) dW, \qquad \hat{F}(0) = f. \tag{2.3}$$

Evaluating the expected values in 2.1a, 2.1b under this model then yields Black's formula,

$$V_{call} = D(t_{set})\{f\mathcal{N}(d_1) - K\mathcal{N}(d_2)\}, \tag{2.4a}$$

$$V_{put} = V_{call} + D(t_{set})[K - f], \tag{2.4b}$$

where

$$d_{1,2} = \frac{\log f/K \pm \frac{1}{2}\sigma_B^2 t_{ex}}{\sigma_B\sqrt{t_{ex}}}, \tag{2.4c}$$

for the price of European calls and puts, as is well-known [11], [12], [13].

All parameters in Black's formula are easily observed, except for the volatility $\sigma_B$..An option's *implied volatility* is the value of $\sigma_B$ that needs to be used in Black's formula so that this formula matches the market price of the option. Since the call (and put) prices in 2.4a - 2.4c are increasing functions of $\sigma_B$, the volatility $\sigma_B$ implied by the market price of an option is unique. Indeed, in many markets it is standard practice to quote prices in terms of the implied volatility $\sigma_B$; the option's dollar price is then recovered by substituting the agreed upon $\sigma_B$ into Black's formula.

The derivation of Black's formula presumes that the volatility $\sigma_B$ is a constant for each underlying asset $\mathcal{A}$. However, the implied volatility needed to match market prices nearly always varies with both the strike $K$ and the time-to-exercise $t_{ex}$. See figure 2.1. Changing the volatility $\sigma_B$ means that a *different* model is being used for the underlying asset for each $K$ and $t_{ex}$. This causes several problems managing large books of options.

The first problem is pricing exotics. Suppose one needs to price a call option with strike $K_1$ which has, say, a down-and-out knock-out at $K_2 < K_1$. Should we use the implied volatility at the call's strike $K_1$, the implied volatility at the barrier $K_2$, or some combination of the two to price this option? Clearly, this option cannot be priced without a single, self-consistent, model that works for all strikes without "adjustments."

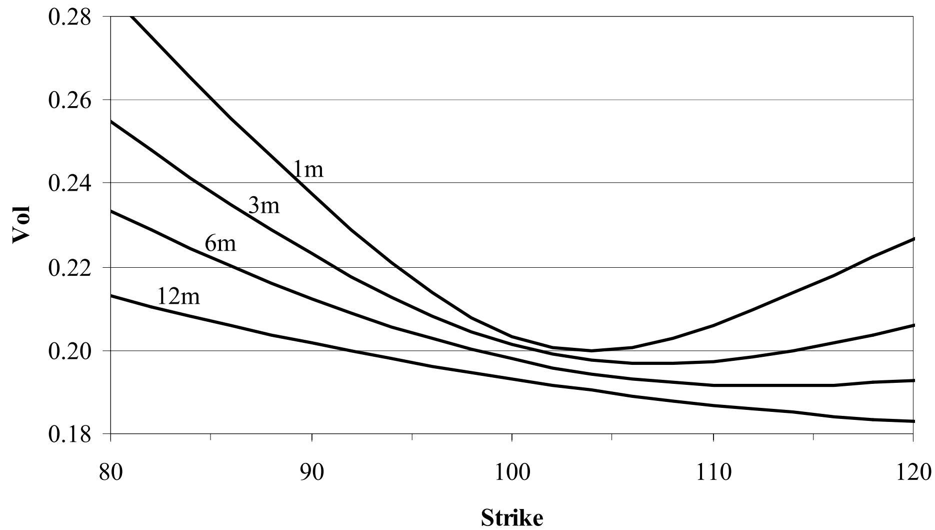


FIG. 2.1. *Implied volatility $\sigma_B(K)$ as a function of the strike $K$ for 1 month, 3 month, 6 month, and 12 month European options on an asset with forward price 100.*

The second problem is hedging. Since *different* models are being used for *different* strikes, it is not clear that the delta and vega risks calculated at one strike are consistent with the same risks calculated at other strikes. For example, suppose that our 1 month option book is long high strike options with a total $\Delta$ risk of $+\$1MM$, and is long low strike options with a $\Delta$ of $-\$1MM$. Is our is our option book really $\Delta$-neutral, or do we have residual delta risk that needs to be hedged? Since different models are used at each strike, it is not clear that the risks offset each other. Consolidating vega risk raises similar concerns. Should we assume parallel or proportional shifts in volatility to calculate the total vega risk of our book? More explicitly, suppose that $\sigma_B$ is 20% at $K = 100$ and 24% at $K = 90$, as shown for the 1m options in figure 2.1. Should we calculate vega by bumping $\sigma_B$ by, say, 0.2% for both options? Or by bumping $\sigma_B$ by 0.2% for the first option and by 0.24% for the second option? These questions are critical to effective book management, since this requires consolidating the delta and vega risks of *all* options on a given asset before hedging, so that only the net exposure of the book is hedged. Clearly one cannot answer these questions without a model that works for all strikes $K$.

The third problem concerns evolution of the implied volatility curve $\sigma_B(K)$. Since the implied volatility

$\sigma_B$ depends on the strike $K$, it is likely to also depend on the current value $f$ of the forward price: $\sigma_B = \sigma_B(f, K)$. In this case there would be systematic changes in $\sigma_B$ as the forward price $f$ of the underlying changes See figure 2.1. Some of the vega risks of Black's model would actually be due to changes in the price of the underlying asset, and should be hedged more properly (and cheaply) as delta risks.

**2.2. Local volatility models.** An apparent solution to these problems is provided by the local volatility model of Dupire [2], which is also attributed to Derman [4], [5]. In an insightful work, Dupire essentially argued that Black was too bold in setting the coefficient $C(t,*)$ to $\sigma_B \hat{F}$. Instead one should only assume that $C$ is Markovian: $C = C(t, \hat{F})$. Re-writing $C(t, \hat{F})$ as $\sigma_{loc}(t, \hat{F})\hat{F}$ then yields the "local volatility model," where the forward price of the asset is

$$d\hat{F} = \sigma_{loc}(t, \hat{F})\hat{F}dW, \qquad \hat{F}(0) = f, \tag{2.5a}$$

in the forward measure. Dupire argued that instead of theorizing about the unknown local volatility function $\sigma_{loc}(t, \hat{F})$, one should obtain $\sigma_{loc}(t, \hat{F})$ directly from the marketplace by "calibrating" the local volatility model to market prices of liquid European options.

In calibration, one starts with a given local volatility function $\sigma_{loc}(t, \hat{F})$, and evaluates

$$V_{call} = D(t_{set})E\left\{[\hat{F}(t_{ex}) - K]^+ | \hat{F}(0) = f, \right\} \tag{2.5b}$$

$$\equiv V_{put} + D(t_{set})(f - K) \tag{2.5c}$$

to obtain the theoretical prices of the options; one then varies the local volatility function $\sigma_{loc}(t, \hat{F})$ until these theoretical prices match the actual market prices of the option for each strike $K$ and exercise date $t_{ex}$. In practice liquid markets usually exist only for options with specific exercise dates $t^1_{ex}, t^2_{ex}, t^3_{ex}, \ldots$; for example, for 1m, 2m, 3m, 6m, and 12m from today. Commonly the local vols $\sigma_{loc}(t, \hat{F})$ are taken to be piecewise constant in time:

$$\begin{aligned} &\sigma_{loc}(t, \hat{F}) = \sigma^{(1)}_{loc}(\hat{F}) && \text{for } t < t^1_{ez}, \\ &\sigma_{loc}(t, \hat{F}) = \sigma^{(j)}_{loc}(\hat{F}) && \text{for } t^{j-1}_{ex} < t < t^j_{ez} \qquad j = 2, 3, ...J \\ &\sigma_{loc}(t, \hat{F}) = \sigma^{(J)}_{loc}(\hat{F}) && \text{for } t > t^J_{ez} \end{aligned} \tag{2.6}$$

One first calibrates $\sigma^{(1)}_{loc}(\hat{F})$ to reproduce the option prices at $t^1_{ex}$ for all strikes $K$, then calibrates $\sigma^{(2)}_{loc}(\hat{F})$ to reproduce the option prices at $t^2_{ex}$, for all $K$, and so forth . This calibration process can be greatly simplified by using the results in [14] and [15]. There we solve to obtain the prices of European options under the local volatility model 2.5a - 2.5c, and from these prices we obtain explicit algebraic formulas for the implied volatility of the local vol models.

Once $\sigma_{loc}(t, \hat{F})$ has been obtained by calibration, the local volatility model is a single, self-consistent model which correctly reproduces the market prices of calls (and puts) for all strikes $K$ and exercise dates $t_{ex}$ without "adjustment." Prices of exotic options can now be calculated from this model without ambiguity. This model yields consistent delta and vega risks for all options, so these risks can be consolidated across strikes. Finally, perturbing $f$ and re-calculating the option prices enables one to determine how the implied volatilites change with changes in the underlying asset price. Thus, the local volatility model provides *a* method of pricing and hedging options in the presence of market smiles and skews. It is perhaps the most popular method of managing exotic equity and foreign exchange options. Unfortunately, the local volatility model predicts the *wrong dynamics* of the implied volatility curve, which leads to inaccruate and often unstable hedges.

To illustrate the problem, consider the special case in which the local vol is a function of $\hat{F}$ only:

$$d\hat{F} = \sigma_{loc}(\hat{F})\hat{F}dW, \qquad \hat{F}(0) = f. \tag{2.7}$$

In [14] and [15] singular perturbation methods were used to analyze this model. There it was found that European call and put prices are given by Black's formula 2.4a - 2.4c with the implied volatility

$$\sigma_B(K,f) = \sigma_{loc}(\tfrac{1}{2}[f+K])\left\{1 + \tfrac{1}{24}\frac{\sigma''_{loc}(\tfrac{1}{2}[f+K])}{\sigma_{loc}(\tfrac{1}{2}[f+K])}(f-K)^2 + \cdots\right. . \tag{2.8}$$

On the right hand side, the first term dominates the solution and the second term provides a much smaller correction The omitted terms are very small, usually less than 1% of the first term.

The behavior of local volatility models can be largely understood by examining the first term in 2.8. The implied volatility depends on both the strike $K$ and the current forward price $f$. So supppose that today the forward price is $f_0$ and the implied volatility curve seen in the marketplace is $\sigma_B^0(K)$. Calibrating the model to the market clearly requires choosing the local volatility to be

$$\sigma_{loc}(\hat{F}) = \sigma_B^0(2\hat{F} - f_0)\{1 + \cdots\}. \tag{2.9}$$

Now that the model is calibrated, let us examine its predictions. Suppose that the forward value changes from $f_0$ to some new value $f$. From 2.8, 2.9 we see that the model predicts that the new implied volatility curve is

$$\sigma_B(K,f) = \sigma_B^0(K + f - f_0)\{1 + \cdots\} \tag{2.10}$$

for an option with strike $K$, given that the current value of the forward price is $f$. In particular, if the forward price $f_0$ *increases* to $f$, the implied volatility curve moves to the *left*; if $f_0$ *decreases* to $f$, the implied volatility curve moves to the *right*. *Local volatility models predict that the market smile/skew moves in the* opposite *direction as the price of the underlying asset.* This is opposite to typical market behavior, in which smiles and skews move in the *same* direction as the underlying.

To demonstrate the problem concretely, suppose that today's implied volatility is a perfect smile

$$\sigma_B^0(K) = \alpha + \beta[K - f_0]^2 \tag{2.11a}$$

around today's forward price $f_0$. Then equation 2.8 implies that the local volatility is

$$\sigma_{loc}(\hat{F}) = \alpha + 3\beta(\hat{F} - f_0)^2 + \cdots . \tag{2.11b}$$

As the forward price $f$ evolves away from $f_0$ due to normal market fluctuations, equation 2.8 predicts that the implied volatility is

$$\sigma_B(K,f) = \alpha + \beta[K - (\tfrac{3}{2}f_0 - \tfrac{1}{2}f)]^2 + \tfrac{3}{4}\beta(f - f_0)^2 + \cdots . \tag{2.11c}$$

. The implied volatility curve not only moves in the opposite direction as the underlying, but the curve also shifts upward regardless of whether $f$ increases or decreases. Exact results are illustrated in figures 2.2 - 2.4. There we assumed that the local volatility $\sigma_{loc}(\hat{F})$ was given by 2.11b, and used finite difference methods to obtain essentially exact values for the option prices, and thus implied volatilites.

Hedges calculated from the local volatility model are wrong. To see this, let $BS(f, K, \sigma_B, t_{ex})$ be Black's formula 2.4a - 2.4c for, say, a call option. Under the local volatility model, the value of a call option is given by Black's formula

$$V_{call} = BS(f, K, \sigma_B(K,f), t_{ex}) \tag{2.12a}$$

with the volatility $\sigma_B(K,f)$ given by 2.8. Differentiating with respect to $f$ yields the $\Delta$ risk

$$\Delta \equiv \frac{\partial V_{call}}{\partial f} = \frac{\partial BS}{\partial f} + \frac{\partial BS}{\partial \sigma_B}\frac{\partial \sigma_B(K,f)}{\partial f}. \tag{2.12b}$$

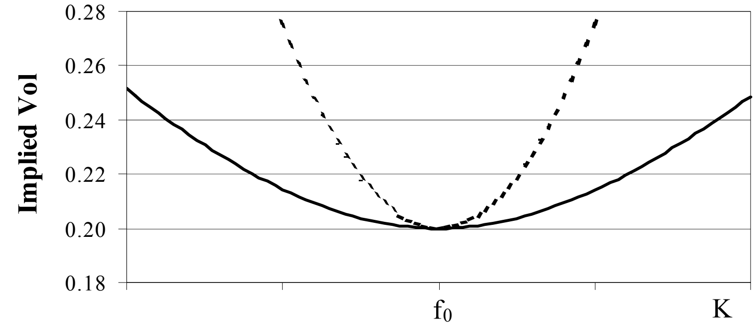


FIG. 2.2. *Exact implied volatility $\sigma_B(K, f_0)$ (solid line) obtained from the local volatility $\sigma_{loc}(\hat{F})$ (dashed line).*

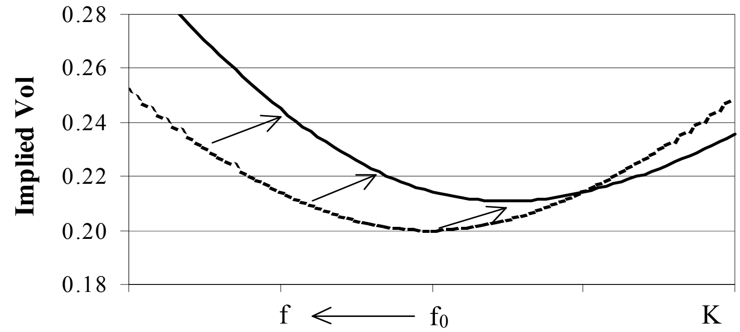


FIG. 2.3. *Implied volatility $\sigma_B(K, f)$ if the forward price* decreases *from $f_0$ to $f$ (solid line).*

predicted by the local volatility model. The first term is clearly the $\Delta$ risk one would calculate from Black's model using the implied volatility from the market. The second term is the local volatility model's correction to the $\Delta$ risk, which consists of the Black vega risk multiplied by the *predicted* change in $\sigma_B$ due to changes in the underlying forward price $f$. In real markets the implied volatily moves in the *opposite* direction as the direction predicted by the model. Therefore, the correction term needed for real markets should have the *opposite sign* as the correction predicted by the local volatility model. *The original Black model yields more accurate hedges than the local volatility model, even though the local vol model is self-consistent across strikes and Black's model is inconsistent.*

Local volatility models are also peculiar theoretically. Using any function for the local volatility $\sigma_{loc}(t, \hat{F})$ *except* for a power law,

$$C(t, *) = \alpha(t)\hat{F}^{\beta}, \tag{2.13}$$

$$\sigma_{loc}(t, \hat{F}) = \alpha(t)\hat{F}^{\beta}/\hat{F} = \alpha(t)/\hat{F}^{1-\beta}, \tag{2.14}$$

introduces an intrinsic "length scale" for the forward price $\hat{F}$ into the model. That is, the model becomes inhomogeneous in the forward price $\hat{F}$. Although intrinsic length scales are theoretically possible, it is difficult to understand the financial origin and meaning of these scales [16], and one naturally wonders whether such

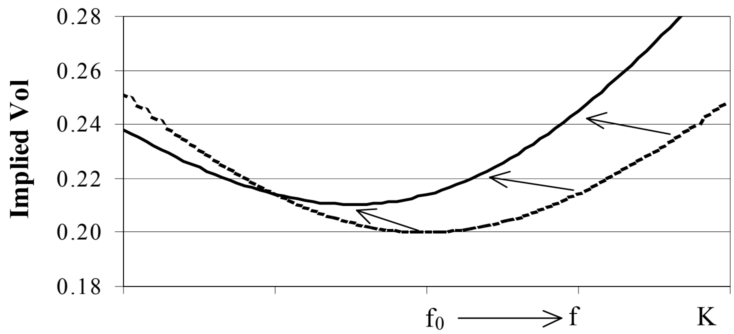


FIG. 2.4. *Implied volatility* $\sigma_B(K, f)$ *if the forward prices* increases *from* $f_0$ *to* $f$ *(solid line).*

scales should be introduced into a model without specific theoretical justification.

**2.3. The SABR model.** The failure of the local volatility model means that we cannot use a Markovian model based on a single Brownian motion to manage our smile risk. Instead of making the model non-Markovian, or basing it on non-Brownian motion, we choose to develop a two factor model. To select the second factor, we note that most markets experience both relatively quiescent and relatively chaotic periods. This suggests that volatility is not constant, but is itself a random function of time. Respecting the preceding discusion, we choose the unknown coefficient $C(t, *)$ to be $\hat{\alpha}\hat{F}^{\beta}$, where the "volatility" $\hat{\alpha}$ is itself a stochastic process. Choosing the simplest reasonable process for $\hat{\alpha}$ now yields the "stochastic-$\alpha\beta\rho$ model," which has become known as the *SABR model.* In this model, the forward price and volatility are

$$d\hat{F} = \hat{\alpha}\hat{F}^{\beta}\, dW_1, \qquad \hat{F}(0) = f \tag{2.15a}$$

$$d\hat{\alpha} = \nu\hat{\alpha}\, dW_2, \qquad \hat{\alpha}(0) = \alpha \tag{2.15b}$$

under the forward measure, where the two processes are correlated by:

$$dW_1\, dW_2 = \rho dt. \tag{2.15c}$$

Many other stochastic volatility models have been proposed, for example [17], [18], [19], [20]; these models will be treated in section 5. However, the SABR model has the virtue of being the simplest stochastic volatility model which is homogenous in $\hat{F}$ and $\hat{\alpha}$. We shall find that the SABR model can be used to accurately fit the implied volatility curves observed in the marketplace for any single exercise date $t_{ex}$. More importantly, it predicts the correct dynamics of the implied volatility curves. This makes the SABR model an effective means to manage the smile risk in markets where each asset only has a single exercise date; these markets include the swaption and caplet/floorlet markets.

As written, the SABR model may or may not fit the observed *volatility surface* of an asset which has European options at several different exercise dates; such markets include foreign exchange options and most equity options. Fitting volatility surfaces requires the *dynamic SABR model* which is introduced and analyzed in section 4.

It has been claimed by many authors that stochastic volatility models are models of incomplete markets, because the stochastic volatility risk cannot be hedged. This is not true. It is true that the risk to changes in $\hat{\alpha}$ (the vega risk) cannot be hedged by buying or selling the underlying asset. However, vega risk can be hedged by buying or selling options on the asset in exactly the same way that $\Delta$-hedging is used to neutralize the risks to changes in the price $\hat{F}$. In practice, vega risks are hedged by buying and selling options as a matter of routine, so whether the market would be complete if these risks were not hedged is a moot question.

The SABR model 2.15a - 2.15c is analyzed in Appendix B. There singular perturbation techniques are used to obtain the prices of European options. From these prices, the options' implied volatility $\sigma_B(K,f)$ is then obtained. The upshot of this analysis is that under the SABR model, the price of European options is given by Black's formula,

$$V_{call} = D(t_{set})\{f\mathcal{N}(d_1) - K\mathcal{N}(d_2)\}, \tag{2.16a}$$

$$V_{put} = V_{call} + D(t_{set})[K - .f], \tag{2.16b}$$

with

$$d_{1,2} = \frac{\log f/K \pm \frac{1}{2}\sigma_B^2 t_{ex}}{\sigma_B\sqrt{t_{ex}}}, \tag{2.16c}$$

where the implied volatility $\sigma_B(f,K)$ is given by

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

Here

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

and $x(z)$ is defined by

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

For the special case of at-the-money options, options struck at $K = f$, this formula reduces to

$$\sigma_{ATM} = \sigma_B(f,f) = \frac{\alpha}{f^{(1-\beta)}}\left\{1 + \left[\frac{(1-\beta)^2}{24}\frac{\alpha^2}{f^{2-2\beta}} + \frac{1}{4}\frac{\rho\beta\alpha\nu}{f^{(1-\beta)}} + \frac{2-3\rho^2}{24}\nu^2\right]t_{ex} + \cdots\right. . \tag{2.18}$$

These formulas are the main result of this paper. Although it appears formidable, the formula is explicit and only involves elementary trignometric functions. Implementing the SABR model for vanilla options is very easy, since once this formula is programmed, we just need to send the options to a Black pricer.In the next section we examine the qualitative behavior of this formula, and how it can be used to managing smile risk.

The complexity of the formula is needed for accurate pricing. Omitting the last line of 2.17a, for example, can result in a relative error that exceeds three per cent in extreme cases. Although this error term seems small, it is large enough to be required for accurate pricing. The omitted terms "$+\cdots$" are *much, much* smaller. Indeed, even though we have derived more accurate expressions by continuing the perturbation expansion to higher order, 2.17a - 2.17c is the formula we use to value and hedge our vanilla swaptions, caps, and floors. We have not implemented the higher order results, believing that the increased precision of the higher order results is superfluous.

There are two special cases of note: $\beta = 1$, representing a stochastic log normal model), and $\beta = 0$, representing a stochastic normal model. The implied volatility for these special cases is obtained in the last section of Appendix B.

**3. Managing smile risk.** The complexity of the above formula for $\sigma_B(K,f)$ obscures the qualitative behavior of the SABR model. To make the model's phenomenology and dynamics more transparent, note that formula 2.17a - 2.17c can be approximated as

$$\text{(3.1a)} \qquad \sigma_B(K,f) = \frac{\alpha}{f^{1-\beta}}\left\{1 - \tfrac{1}{2}(1-\beta-\rho\lambda)\log K/f \right. \\ \left. + \tfrac{1}{12}\left[(1-\beta)^2 + (2-3\rho^2)\lambda^2\right]\log^2 K/f + \cdots\right.,$$

provided that the strike $K$ is not too far from the current forward $f$. Here the ratio

$$\text{(3.1b)} \qquad \lambda = \frac{\nu}{\alpha} f^{1-\beta}$$

measures the strength $\nu$ of the volatility of volatility (the "volvol") compared to the local volatility $\alpha/f^{1-\beta}$ at the current forward. Although equations 3.1a - 3.1b should not be used to price real deals, they are accurate enough to depict the qualitative behavior of the SABR model faithfully.

As $f$ varies during normal trading, the curve that the ATM volatility $\sigma_B(f,f)$ traces is known as the *backbone*, while the *smile* and *skew* refer to the implied volatility $\sigma_B(K,f)$ as a function of strike $K$ for a fixed $f$. That is, the market smile/skew gives a snapshot of the market prices for different strikes $K$ at a given instance, when the forward $f$ has a specific price. Figures 3.1 and 3.2. show the dynamics of the smile/skew predicted by the SABR model.

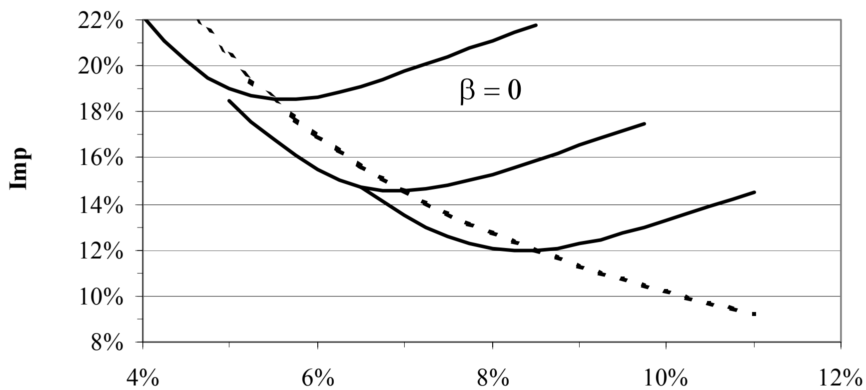


FIG. 3.1. *Backbone and smiles for $\beta = 0$. As the forward $f$ varies, the implied volatiliity $\sigma_B(f,f)$ of ATM options traverses the backbone (dashed curve). Shown are the smiles $\sigma_B(K,f)$ for three different values of the forward. Volatility data from 1 into 1 swaption on 4/28/00, courtesy of Cantor-Fitzgerald.*

Let us now consider the implied volatility $\sigma_B(K,f)$ in detail. The first factor $\alpha/f^{1-\beta}$ in 3.1a is the implied volatility for at-the-money (ATM) options, options whose strike $K$ equals the current forward $f$. So the backbone traversed by ATM options is essentially $\sigma_B(f,f) = \alpha/f^{1-\beta}$ for the SABR model. The backbone is almost entirely determined by the exponent $\beta$, with the exponent $\beta = 0$ (a stochastic Gaussian model) giving a steeply downward sloping backbone, and the exponent $\beta = 1$ giving a nearly flat backbone.

The second term $-\frac{1}{2}(1-\beta-\rho\lambda)\log K/f$ represents the skew, the slope of the implied volatility with respect to the strike $K$. The $-\frac{1}{2}(1-\beta)\log K/f$ part is the *beta skew*, which is downward sloping since $0 \le \beta \le 1$. It arises because the "local volatility" $\hat{\alpha}\hat{F}^{\beta}/\hat{F}^{1} = \hat{\alpha}/\hat{F}^{1-\beta}$ is a decreasing function of the forward price. The second part $\frac{1}{2}\rho\lambda\log K/f$ is the *vanna* skew, the skew caused by the correlation between

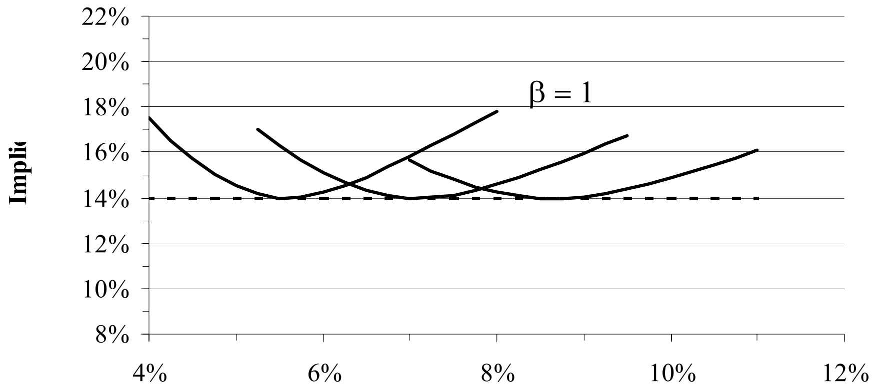


FIG. 3.2. *Backbone and smiles as above, but for $\beta = 1$.*

the volatility and the asset price. Typically the volatility and asset price are negatively correlated, so on average, the volatility $\alpha$ would decrease (increase) when the forward $f$ increases (decreases). It thus seems unsurprising that a negative correlation $\rho$ causes a downward sloping vanna skew.

It is interesting to compare the skew to the slope of the backbone. As $f$ changes to $f'$ the ATM vol changes to

$$\sigma_B(f', f') = \frac{\alpha}{f^{1-\beta}}\{1 - (1-\beta)\frac{f'-f}{f} + \cdots\}. \tag{3.2a}$$

Near $K = f$, the $\beta$ component of skew expands as

$$\sigma_B(K, f) = \frac{\alpha}{f^{1-\beta}}\{1 - \tfrac{1}{2}(1-\beta)\frac{K-f}{f} + \cdots\}, \tag{3.2b}$$

so the slope of the backbone $\sigma_B(f, f)$ is twice as steep as the skew in $\sigma_B(K, f)$ caused by $\beta$.

The last term in 3.1a also contains two parts. The first part $\frac{1}{12}(1-\beta)^2 \log^2 K/f$ appears to be a smile (quadratic) term, but it is dominated by the downward sloping beta skew, and, at reasonable strikes $K$, it just modifies this skew somewhat. The second part $\frac{1}{12}(2 - 3\rho^2)\lambda^2 \log^2 K/f$ is the smile induced by the *volga* (vol-gamma) effect. Physically this smile arises because of "adverse selection": unusually large movements of the forward $\hat{F}$ happen more often when the volatility $\alpha$ increases, and less often when $\alpha$ decreases, so strikes $K$ far from the money represent, on average, high volatility environments.

**3.1. Fitting market data.** The exponent $\beta$ and correlation $\rho$ affect the volatility smile in similar ways. They both cause a downward sloping skew in $\sigma_B(K, f)$ as the strike $K$ varies. From a single market snapshot of $\sigma_B(K, f)$ as a function of $K$ at a given $f$, it is difficult to distinguish between the two parameters. This is demonstrated by figure 3.3. There we fit the SABR parameters $\alpha, \rho, \nu$ with $\beta = 0$ and then re-fit the parameters $\alpha, \rho, \nu$ with $\beta = 1$. Note that there is no substantial difference in the quality of the fits, despite the presence of market noise. This matches our general experience: market smiles can be fit equally well with any specific value of $\beta$. In particular, $\beta$ cannot be determined by fitting a market smile since this would clearly amount to "fitting the noise."

Figure 3.3 also exhibits a common data quality issue. Options with strikes $K$ away from the current forward $f$ trade less frequently than at-the-money and near-the-money options. Consequently, as $K$ moves

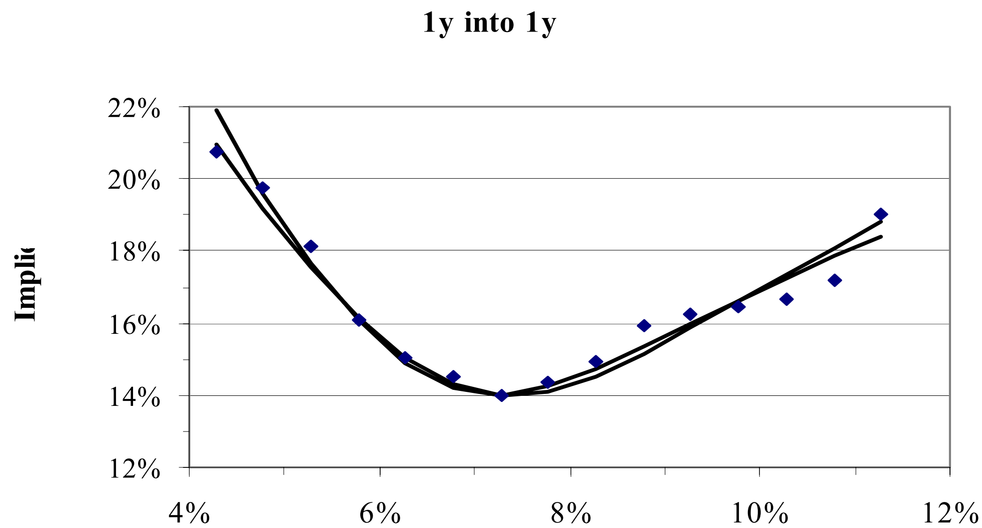


FIG. 3.3. *Implied volatilities as a function of strike. Shown are the curves obtained by fitting the SABR model with exponent $\beta = 0$ and with $\beta = 1$ to the 1y into 1y swaption vol observed on 4/28/00. As usual, both fits are equally good. Data courtesy of Cantor-Fitzgerald.*

away from $f$, the volatility quotes become more suspect because they are more likely to be out-of-date and not represent bona fide offers to buy or sell options.

Suppose for the moment that the exponent $\beta$ is known or has been selected. Taking a snapshot of the market yields the implied volatility $\sigma_B(K, f)$ as a function of the strike $K$ at the current forward price $f$. With $\beta$ given, fitting the SABR model is a straightforward procedure. The three parameters $\alpha, \rho$, and $\nu$ have different effects on the curve: the parameter $\alpha$ mainly controls the overall height of the curve, changing the correlation $\rho$ controls the curve's skew, and changing the vol of vol $\nu$ controls how much smile the curve exhibits. Because of the widely seperated roles these parameters play, the fitted parameter values tend to be very stable, even in the presence of large amounts of market noise.

The exponent $\beta$ can be determined from historical observations of the "backbone" or selected from "aesthetic considerations." Equation 2.18 shows that the implied volatility of ATM options is

$$\log \sigma_B(f, f) = \log \alpha - (1 - \beta) \log f + \log \left\{ 1 + [\tfrac{(1-\beta)^2}{24} \frac{\alpha^2}{f^{2-2\beta}} + \tfrac{1}{4} \frac{\rho \beta \alpha \nu}{f^{(1-\beta)}} + \tfrac{2-3\rho^2}{24} \nu^2] t_{ex} + \cdots . \right\} \tag{3.3}$$

The exponent $\beta$ can be extracted from a log log plot of historical observations of $f, \sigma_{ATM}$ pairs. Since both $f$ and $\alpha$ are stochastic variables, this fitting procedure can be quite noisy, and as the $[\cdots]t_{ex}$ term is typically less than one or two per cent, it is usually ignored in fitting $\beta$.

Selecting $\beta$ from "aesthetic" or other *a priori* considerations usually results in $\beta = 1$ (stochastic lognormal), $\beta = 0$ (stochastic normal), or $\beta = \frac{1}{2}$ (stochastic CIR) models. Proponents of $\beta = 1$ cite log normal models as being "more natural." or believe that the horizontal backbone best represents their market. These proponents often include desks trading foreign exchange options. Proponents of $\beta = 0$ usually believe that a normal model, with its symmetric break-even points, is a more effective tool for managing risks, and would claim that $\beta = 0$ is essential for trading markets like Yen interest rates, where the forwards $f$ can be negative or near zero. Proponents of $\beta = \frac{1}{2}$ are usually US interest rate desks that have developed trust in CIR models.

It is usually more convenient to use the at-the-money volatility $\sigma_{ATM}, \beta, \rho$, and $\nu$ as the SABR parameters instead of the original parameters $\alpha$,$\beta$, $\rho$, $\nu$.The parameter $\alpha$ is then found whenever needed by inverting 2.18 on the fly; this inversion is numerically easy since the $[\cdots]t_{ex}$ term is small. With this parameterization, fitting the SABR model requires fitting $\rho$ and $\nu$ to the implied volatility curve, with $\sigma_{ATM}$ and $\beta$ given. In many markets, the ATM volatilities need to be updated frequently, say once or twice a day, while the smiles and skews need to be updated infrequently, say once or twice a month. With the new parameterization, $\sigma_{ATM}$ can be updated as often as needed, with $\rho$, $\nu$ (and $\beta$) updated only as needed.

Let us apply SABR to options on US dollar interest rates. There are three key groups of European options on US rates: Eurodollar future options, caps/floors, and European swaptions. Eurodollar future options are exchange-traded options on the 3 month Libor rate; like interest rate futures, EDF options are quoted on $100(1-r_{Libor})$. Figure 1.1 fits the SABR model (with $\beta = 1$) to the implied volatility for the June 99 contracts, and figures 3.4 - 3.7 fit the model (also with $\beta = 1$) to the implied volatility for the September 99, December 99, and March 00 contracts. All prices were obtained from Bloomberg Information Services on March 23, 1999. Two points are shown for the same strike where there are quotes for both puts and calls. Note that market liquidity dries up for the later contracts, and for strikes that are too far from the money. Consequently, more market noise is seen for these options.

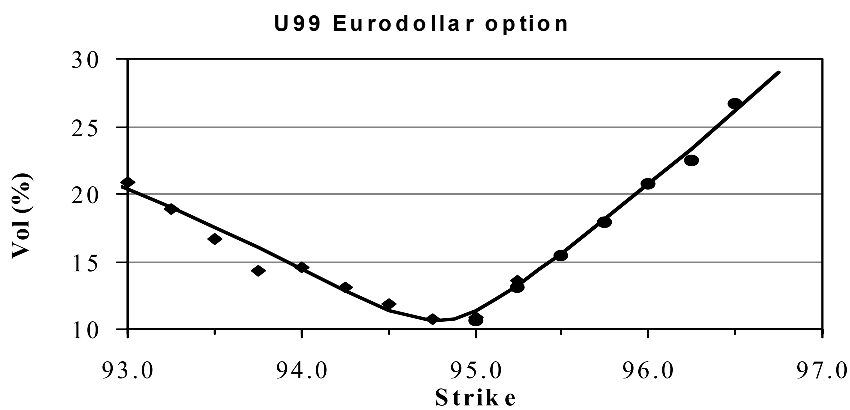


FIG. 3.4. *Volatility of the Sep 99 EDF options*

Caps and floors are sums of caplets and floorlets; each caplet and floorlet is a European option on the 3 month Libor rate. We do not consider the cap/floor market here because the broker-quoted cap prices must be "stripped" to obtain the caplet volatilities before SABR can be applied.

A *m year into n year* swaption is a European option with m years to the exercise date (the maturity); if it is exercised, then one receives an n year swap (the tenor, or underlying) on the 3 month Libor rate. See Appendix A. For almost all maturities and tenors, the US swaption market is liquid for at-the-money swaptions, but is ill-liquid for swaptions struck away from the money. Hence, market data is somewhat suspect for swaptions that are not struck near the money. Figures 3.8 - 3.11 fits the SABR model (with $\beta = 1$) to the prices of $m$ into$5Y$ swaptions observed on April 28, 2000. Data supplied courtesy of Cantor-Fitzgerald.

We observe that the smile and skew depend heavily on the time-to-exercise for Eurodollar future options and swaptions. The smile is pronounced for short-dated options and flattens for longer dated options; the skew is overwhelmed by the smile for short-dated options, but is important for long-dated options. This

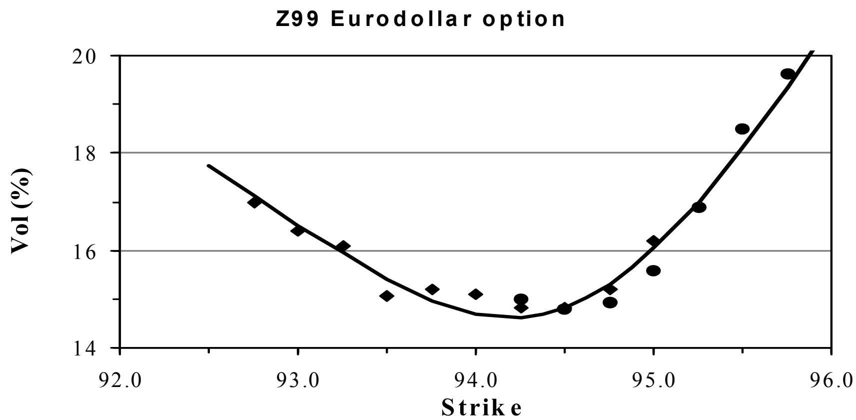


FIG. 3.5. *Volatility of the Dec 99 EDF options*

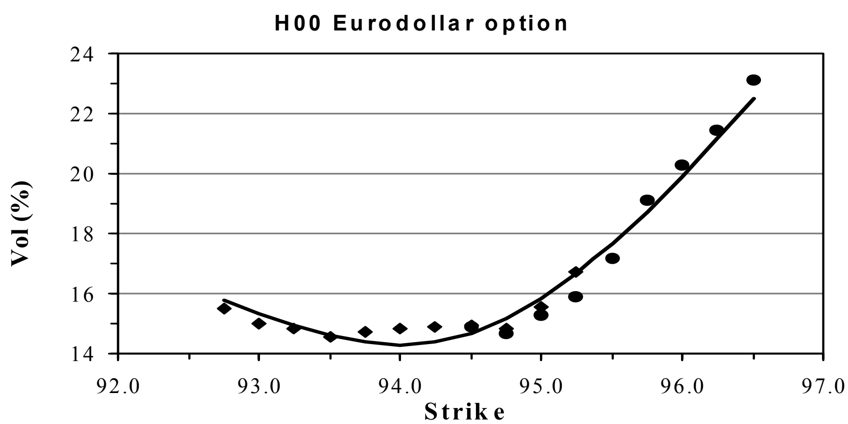


FIG. 3.6. *Volatility of the Mar 00 EDF options*

picture is confirmed tables 3.1 and 3.2. These tables show the values of the vol of vol $\nu$ and correlation $\rho$ obtained by fitting the smile and skew of each "m into n" swaption, again using the data from April 28, 2000. Note that the vol of vol $\nu$ is very high for short dated options, and decreases as the time-to-exercise increases, while the correlations starts near zero and becomes substantially negative. Also note that there is little dependence of the market skew/smile on the length of the underlying swap; both $\nu$ and $\rho$ are fairly constant across each row. This matches our general experience: in most markets there is a strong smile for short-dated options which relaxes as the time-to-expiry increases; consequently the volatility of volatility $\nu$ is large for short dated options and smaller for long-dated options, regardless of the particular underlying. Our experience with correlations is less clear: in some markets a nearly flat skew for short maturity options develops into a strongly downward sloping skew for longer maturities. In other markets there is a strong downward skew for all option maturities, and in still other markets the skew is close to zero for all maturities.

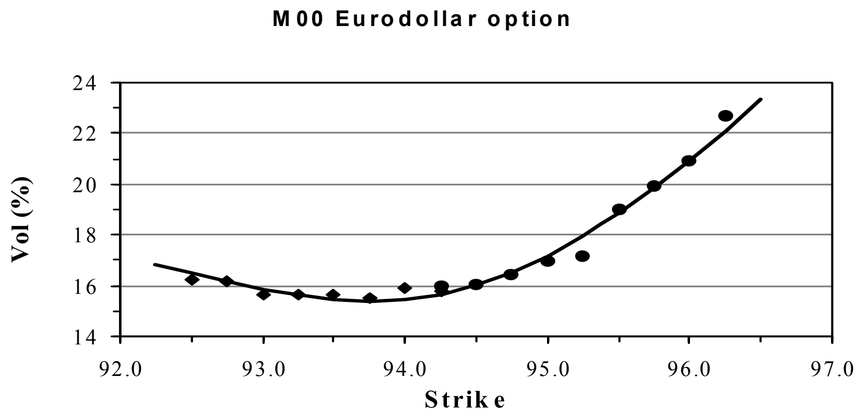


FIG. 3.7. *Volatility of the Jun 00 EDF options*

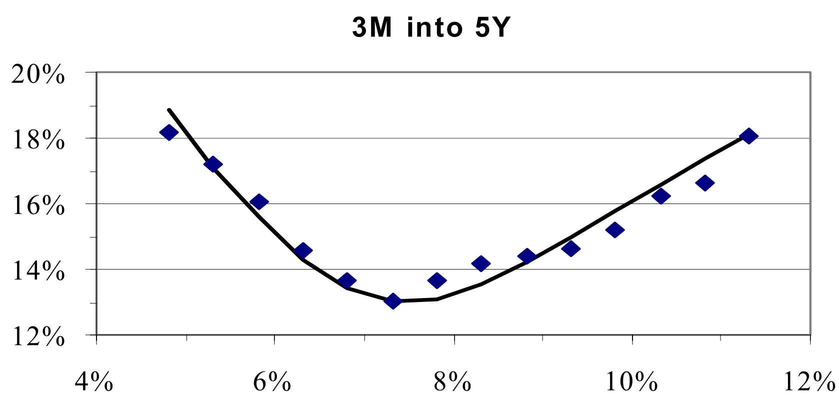


FIG. 3.8. *Volatilities of 3 month into 5 year swaption*

**3.2. Managing smile risk.** After choosing $\beta$ and fitting $\rho$, $\nu$, and either $\alpha$ or $\sigma_{ATM}$, the SABR model

$$d\hat{F} = \hat{\alpha}\hat{F}^{\beta}\, dW_1, \qquad \hat{F}(0) = f \tag{3.4a}$$

$$d\hat{\alpha} = \nu\hat{\alpha}\, dW_2, \qquad \hat{\alpha}(0) = \alpha \tag{3.4b}$$

with

$$dW_1\, dW_2 = \rho dt \tag{3.4c}$$

fits the smiles and skews observed in the market quite well, especially considering the quality of price quotes away from the money . Let us take for granted that it fits well enough. Then we have a single, self-consistent model that fits the option prices for all strikes $K$ without “adjustment,” so we can use this model to price exotic options without ambiguity. The SABR model also predicts that whenever the forward price $f$ changes,

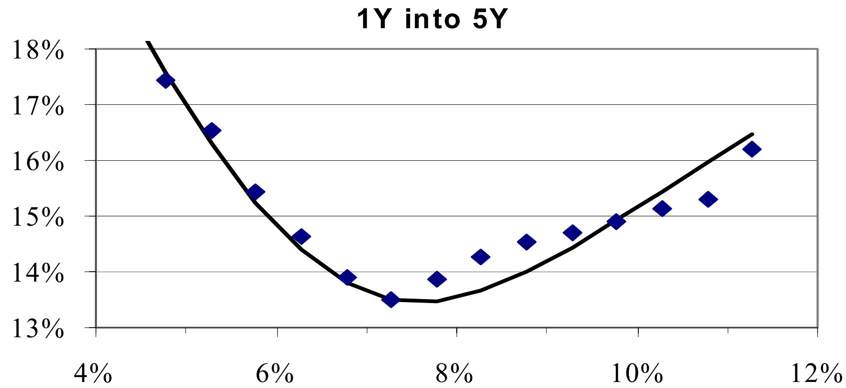


FIG. 3.9. *Volatilities of 1 year into 1 year swaptions*

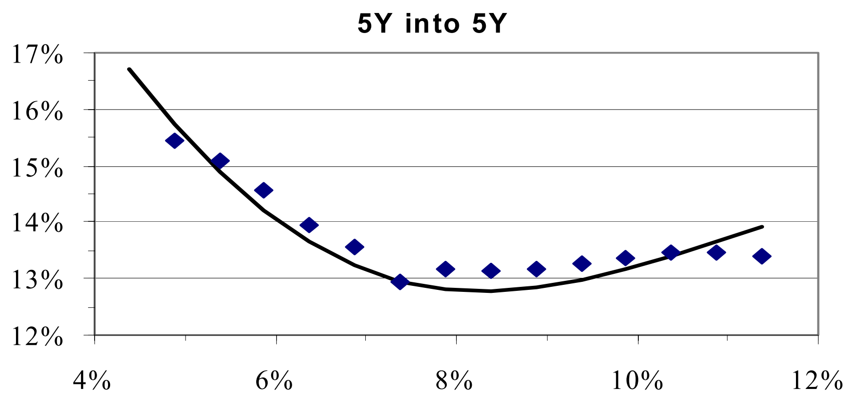


FIG. 3.10. *Volatilities of 5 year into 5 year swaptions*

the the implied volatility curve shifts in the *same* direction and by the *same* amount as the price $f$. This predicted dynamics of the smile matches market experience. If $\beta < 1$, the "backbone" is downward sloping, so the shift in the implied volatility curve is not purely horizontal. Instead, this curve shifts up and down as the at-the-money point traverses the backbone. Our experience suggests that the parameters $\rho$ and $\nu$ are very stable ($\beta$ is assumed to be a given constant), and need to be re-fit only every few weeks. This stability may be because the SABR model reproduces the usual dynamics of smiles and skews. In contrast, the at-the-money volatility $\sigma_{ATM}$, or, equivalently, $\alpha$ may need to be updated every few hours in fast-paced markets.

Since the SABR model is a single self-consistent model for all strikes $K$, the risks calculated at one strike are consistent with the risks calculated at other strikes. Therefore the risks of all the options on the same asset can be added together, and only the residual risk needs to be hedged.

Let us set aside the $\Delta$ risk for the moment, and calculate the other risks. Let $BS(f, K, \sigma_B, t_{ex})$ be

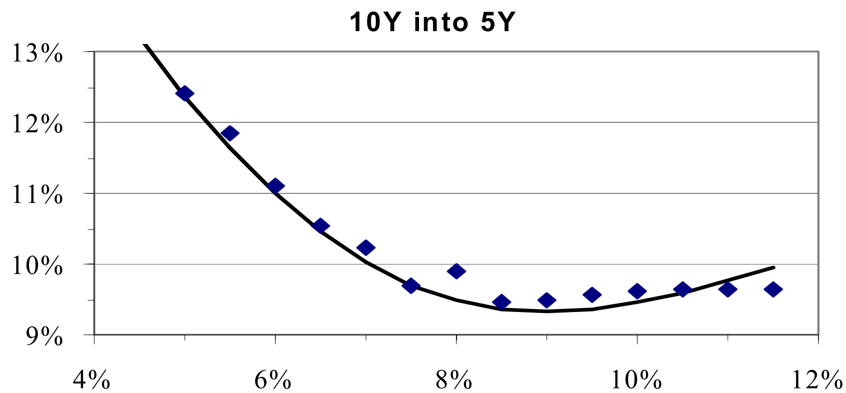


FIG. 3.11. *Volatilities of 10 year into 5 year options*

| $\nu$ | 1Y | 2Y | 3Y | 4Y | 5Y | 7Y | 10Y |
|---|---|---|---|---|---|---|---|
| 1M | 76.2% | 75.4% | 74.6% | 74.1% | 75.2% | 73.7% | 74.1% |
| 3M | 65.1% | 62.0% | 60.7% | 60.1% | 62.9% | 59.7% | 59.5% |
| 6M | 57.1% | 52.6% | 51.4% | 50.8% | 49.4% | 50.4% | 50.0% |
| 1Y | 59.8% | 49.3% | 47.1% | 46.7% | 46.0% | 45.6% | 44.7% |
| 3Y | 42.1% | 39.1% | 38.4% | 38.4% | 36.9% | 38.0% | 37.6% |
| 5Y | 33.4% | 33.2% | 33.1% | 32.6% | 31.3% | 32.3% | 32.2% |
| 7Y | 30.2% | 29.2% | 29.0% | 28.2% | 26.2% | 27.2% | 27.0% |
| 10Y | 26.7% | 26.3% | 26.0% | 25.6% | 24.8% | 24.7% | 24.5% |

TABLE 3.1

*Volatility of volatility $\nu$ for European swaptions. Rows are time-to-exercise; columns are tenor of the underlying swap.*

Black's formula 2.4a - 2.4c for, say, a call option. According to the SABR model, the value of a call is

$$V_{call} = BS(f, K, \sigma_B(K, f), t_{ex}) \tag{3.5}$$

where the volatility $\sigma_B(K, f) \equiv \sigma_B(K, f; \alpha, \beta, \rho, \nu)$ is given by equations 2.17a - 2.17c. Differentiating[1] with respect to $\alpha$ yields the vega risk, the risk to overall changes in volatility:

$$\frac{\partial V_{call}}{\partial \alpha} = \frac{\partial BS}{\partial \sigma_B} \cdot \frac{\partial \sigma_B(K, f; \alpha, \beta, \rho, \nu)}{\partial \alpha}. \tag{3.6}$$

This risk is the change in value when $\alpha$ changes by a unit amount. It is traditional to scale vega so that it represents the change in value when the ATM volatility changes by a unit amount. Since $\delta\sigma_{ATM} = (\partial\sigma_{ATM}/\partial\alpha)\delta\alpha$, the vega risk is

$$\text{vega} \equiv \frac{\partial V_{call}}{\partial \alpha} = \frac{\partial BS}{\partial \sigma_B} \cdot \frac{\dfrac{\partial \sigma_B(K, f; \alpha, \beta, \rho, \nu)}{\partial \alpha}}{\dfrac{\partial \sigma_{ATM}(f; \alpha, \beta, \rho, \nu)}{\partial \alpha}} \tag{3.7a}$$

[1] In practice risks are calculated by finite differences: valuing the option at $\alpha$, re-valuing the option after bumping $\alpha$ to $\alpha + \delta$, and then subtracting to determine the risk. This saves differentiating complex formulas such as 2.17a - 2.17c.

| $\rho$ | 1Y | 2Y | 3Y | 4Y | 5Y | 7Y | 10Y |
|---|---|---|---|---|---|---|---|
| 1M | 4.2% | -0.2% | -0.7% | -1.0% | -2.5% | -1.8% | -2.3% |
| 3M | 2.5% | -4.9% | -5.9% | -6.5% | -6.9% | -7.6% | -8.5% |
| 6M | 5.0% | -3.6% | -4.9% | -5.6% | -7.1% | -7.0% | -8.0% |
| 1Y | -4.4% | -8.1% | -8.8% | -9.3% | -9.8% | -10.2% | -10.9% |
| 3Y | -7.3% | -14.3% | -17.1% | -17.1% | -16.6% | -17.9% | -18.9% |
| 5Y | -11.1% | -17.3% | -18.5% | -18.8% | -19.0% | -20.0% | -21.6% |
| 7Y | -13.7% | -22.0% | -23.6% | -24.0% | -25.0% | -26.1% | -28.7% |
| 10Y | -14.8% | -25.5% | -27.7% | -29.2% | -31.7% | -32.3% | -33.7% |

Table 3.2

*Matrix of correlations $\rho$ between the underlying and the volatility for European swaptons.*

where $\sigma_{ATM}(f) = \sigma_B(f, f)$ is given by 2.18. Note that to leading order, $\partial\sigma_B/\partial\alpha \approx \sigma_B/\alpha$ and $\partial\sigma_{ATM}/\partial\alpha \approx \sigma_{ATM}/\alpha$, so the vega risk is roughly given by

$$\text{vega} \approx \frac{\partial BS}{\partial\sigma_B} \cdot \frac{\sigma_B(K,f)}{\sigma_{ATM}(f)} = \frac{\partial BS}{\partial\sigma_B} \cdot \frac{\sigma_B(K,f)}{\sigma_B(f,f)}. \tag{3.7b}$$

Qualitatively, then, vega risks at different strikes are calculated by bumping the implied volatility at each strike $K$ by an amount that is proportional to the implied volatiity $\sigma_B(K, f)$ at that strike. That is, in using equation 3.7a, we are essentially using proportional, and not parallel, shifts of the volatility curve to calculate the total vega risk of a book of options.

Since $\rho$ and $\nu$ are determined by fitting the implied volatility curve observed in the marketplace, the SABR model has risks to $\rho$ and $\nu$ changing. Borrowing terminology from foreign exchange desks, vanna is the risk to $\rho$ changing and volga (vol gamma) is the risk to $\nu$ changing:

$$\text{vanna} = \frac{\partial V_{call}}{\partial\rho} = \frac{\partial BS}{\partial\sigma_B} \cdot \frac{\partial\sigma_B(K, f; \alpha, \beta, \rho, \nu)}{\partial\rho}, \tag{3.8a}$$

$$\text{volga} = \frac{\partial V_{call}}{\partial\nu} = \frac{\partial BS}{\partial\sigma_B} \cdot \frac{\partial\sigma_B(K, f; \alpha, \beta, \rho, \nu)}{\partial\nu}. \tag{3.8b}$$

Vanna basically expresses the risk to the skew increasing, and volga expresses the risk to the smile becoming more pronounced. These risks are easily calculated by using finite differences on the formula for $\sigma_B$ in equations 2.17a - 2.17c. If desired, these risks can be hedged by buying or selling away-from-the-money options.

The delta risk expressed by the SABR model depends on whether one uses the parameterization $\alpha$, $\beta$, $\rho$, $\nu$ or $\sigma_{ATM}$, $\beta$, $\rho$, $\nu$. Suppose first we use the parameterization $\alpha$, $\beta$, $\rho$, $\nu$, so that $\sigma_B(K, f) \equiv \sigma_B(K, f; \alpha, \beta, \rho, \nu)$. Differentiating respect to $f$ yields the $\Delta$ risk

$$\Delta \equiv \frac{\partial V_{call}}{\partial f} = \frac{\partial BS}{\partial f} + \frac{\partial BS}{\partial\sigma_B}\frac{\partial\sigma_B(K, f; \alpha, \beta, \rho, \nu)}{\partial f}. \tag{3.9}$$

The first term is the ordinary $\Delta$ risk one would calculate from Black's model. The second term is the SABR model's correction to the $\Delta$ risk. It consists of the Black vega times the *predicted* change in the implied volatility $\sigma_B$ caused by the change in the forward $f$. As discussed above, the predicted change consists of a sideways movement of the volatility curve in the same direction (and by the same amount) as the change in the forward price $f$. In addition, if $\beta < 1$ the volatility curve rises and falls as the at-the-money point traverses up and down the backbone. There may also be minor changes to the shape of the skew/smile due to changes in $f$.

Now suppose we use the parameterization $\sigma_{AMT}$, $\beta$, $\rho$, $\nu$. Then $\alpha$ is a function of $\sigma_{ATM}$ and $f$ defined implicitly by 2.18. Differentiating 3.5 now yields the $\Delta$ risk

$$\Delta \equiv \frac{\partial BS}{\partial f} + \frac{\partial BS}{\partial \sigma_B}\left\{\frac{\partial \sigma_B(K,f;\alpha,\beta,\rho,\nu)}{\partial f} + \frac{\partial \sigma_B(K,f;\alpha,\beta,\rho,\nu)}{\partial \alpha}\frac{\partial \alpha(\sigma_{ATM},f)}{\partial f}\right\}. \tag{3.10}$$

The delta risk is now the risk to changes in $f$ with $\sigma_{ATM}$ held fixed. The last term is just the change in $\alpha$ needed to keep $\sigma_{ATM}$ constant while $f$ changes. Clearly this last term must just cancel out the vertical component of the backbone, leaving only the sideways movement of the implied volatilty curve. Note that this term is zero for $\beta = 1$.

Theoretically one should use the $\Delta$ from equation 3.9 to risk manage option books. In many markets, however, it may take several days for volatilities $\sigma_B$ to change following significant changes in the forward price $f$. In these markets, using $\Delta$ from 3.10 is a *much* more effective hedge. For suppose one used $\Delta$ from equation 3.9. Then, when the volatility $\sigma_{ATM}$ did not immediately change following a change in $f$, one would be forced to re-mark $\alpha$ to compensate, and this re-marking would change the $\Delta$ hedges. As $\sigma_{ATM}$ equilibrated over the next few days, one would mark $\alpha$ back to its original value, which would change the $\Delta$ hedges back to their original value. This "hedging chatter" caused by market delays can prove to be costly.

**4. The dynamic SABR model.** Quote results for smile and skew.
For each exercise date, same smile as in the static SABR model! Same smile dynamics!
Calibrating volatility surface is no harder than calibrating smile.
Show some results. FX options?

**Appendix A. Martingale pricing.**

Martingale pricing theory [7], [6], [8], [27] envisions an economy in which all asset prices can be generated by a finite set of Brownian motions; in which all asset prices are non-anticipating; and in which all asset prices are regular enough to be written as Ito processes. Along with some technical assumptions (the "usual conditions"), these postulates lead to the fundamental theorem of arbitrage free finance:

Theorem A.1. *Let the numeraire N(t) be the value of any freely tradeable, positively valued asset which generates no cash flow. In the absence of arbitrage, there exists a probability measure Q (which depends on the security chosen as the numeraire) such that for all other securities, the value V(t) of the security divided by the numeraire N(t) is a Martingale. Specifically, if V(t) generates no cash flow, then at any time t,*

$$\frac{V(t)}{N(t)} = E^Q\left\{\frac{V(T)}{N(T)}|\mathfrak{F}_t\right\} \qquad \text{for all } T > t. \tag{A.1a}$$

*where the expected value $E^Q$ is over this probability measure. If V(t) generates cash flow C(t), then*

$$\frac{V(t)}{N(t)} = E^Q\left\{\frac{V(T)}{N(T)} + \int_t^T \frac{C(T')}{N(T')}dT' \;|\mathfrak{F}_t\right\} \qquad \text{for all } T > t. \tag{A.1b}$$

In this appendix, we apply use this theorem to develop the theoretical framework for pricing European options and swaptions.

**A.1. General European options.** Consider a financial asset $\mathcal{A}$. Suppose we have a forward contact on this asset which requires us to pay the *strike price* $K$ and acquire the asset on the *settlement date* $t_{set}$. Let $\hat{A}(t)$ be the asset's current price, let $V_{fc}(t)$ be the value of the forward contract, and let $Z(t;t_{set})$ be the value of a zero coupon bond which pays \$1 on the settlement date $t_{set}$.

To price the forward contract, let us choose the zero coupon bond for our numeraire. Then there exists a probability measure $Q$ (known as the forward measure), such that $V_{fc}(t)/Z(t;t_{set})$ is a Martingale:

$$\text{(A.2)} \qquad \frac{V_{fc}(t)}{Z(t,t_{set})} = E^Q\left\{\frac{V_{fc}(T)}{Z(T;t_{set})}|\mathfrak{F}_t\right\} \qquad \text{for all } T > t$$

At the settlement date, clearly $V_{fc}(t_{set}) = \hat{A}(t_{set}) - K$ and $Z(t_{set};t_{set}) = 1$, so evaluating this expression at $T = t_{set}$ yields

$$\text{(A.3)} \qquad V_{fc}(t) = Z(t;t_{set})\left[E^Q\left\{\hat{A}(t_{set})|\mathfrak{F}_t\right\} - K\right].$$

Clearly

$$\text{(A.4)} \qquad \hat{F}(t) = E^Q\left\{\hat{A}(t_{set})|\mathfrak{F}_t\right\}$$

is the *forward price* of the asset for date $t_{set}$ as seen at date $t$, since it is the strike at which the forward contract would have no value on date $t$. The value of the forward contract can now be written as

$$\text{(A.5)} \qquad V_{fc}(t) = Z(t;t_{set})\left[\hat{F}(t) - K\right],$$

the usual formula for a forward contract.

Observe that the forward price is a Martingale in the forward measure; this can be seen from the "telescoping" rule of expected values. Alternatively, one can simply note that since $Z(t;t_{set})$ is our numeraire, and a forward contract is certainly a tradeable security, the fundamental theorem shows that $V_{fc}(t)/Z(t;t_{set})$, and thus $\hat{F}(t) - K$, is a Martingale. Since $K$ is a constant, clearly $\hat{F}(t)$ is a Martingale. The Martingale representation theorem now shows that

$$\text{(A.6)} \qquad d\hat{F} = C(t,*)dW,$$

for some coefficient $C(t,*)$, where $dW$ is Brownian motion under $Q$.

Now consider a European call option on the same asset $\mathcal{A}$.If one exercises a European option on its exercise date $t_{ex}$, then the asset will be acquired, and the strike paid, on the settlement date $t_{set}$ of the option. That is, when one exercises a European option on $t_{ex}$, one usually receives a forward contract on the asset, and not immediate possesion of the asset.

To price the European call option, let us again choose the zero coupon bond $Z(t;t_{set})$ as the numeraire. The value of the option is

$$\text{(A.7)} \qquad \frac{V_{call}(t)}{Z(t,t_{set})} = E^Q\left\{\frac{V_{call}(T)}{Z(T;t_{set})}|\mathfrak{F}_t\right\} \qquad \text{for all } T > t.$$

The simplest date $T$ to use is the exercise date $t_{ex}$. Since the option will be exercised if the forward contract has positive value, the value of the option at date $T = t_{ex}$ is clearly $[V_{fc}(t_{ex}]^+$, which is $Z(t_{ex};t_{set})\left[\hat{F}(t_{ex}) - K\right]^+$. Substituting this into the above equation yields

$$\text{(A.8a)} \qquad V_{call}(t) = Z(t,t_{set})E^Q\left\{\left[\hat{F}(t_{ex}) - K\right]^+|\mathfrak{F}_t\right\}$$

for the value of the European option, where the expectation is over the forward measure. Moreover, in this measure, the forward price is a Martingale,

$$\text{(A.8b)} \qquad dF = C(t,*)dW.$$

Letting today be $t = 0$, and noting that today's values of zero coupon bonds $Z(0,t_{set})$ are just the discount factors $D(t_{set})$, now reduces A.8a, A.8b to equations 2.1a, 2.1b, the starting point of our analysis..

**A.2. European swaptions.** European swaptions fit naturally into the same framework. A standard swap is defined by specifying the fixed rate (strike) $R_f$, the start date $t_0$, and the theoretical end date $\tilde{t}_n$. In most currencies, the default frequency for the fixed leg is semiannual, and the *theoretical dates* $\tilde{t}_1, \tilde{t}_2, \ldots, \tilde{t}_n$ are generated by subtracting 6 months, 12 months, 18 months, ... from the theoretical end date $\tilde{t}_n$. The *actual dates* $t_1, t_2, \ldots, t_n$ are then obtained by replacing any bad business days according to the *modified following* business day convention. Once the actual dates are constructed, the fixed leg pays

$$\alpha_j R_f \qquad \text{paid at } t_j \qquad \text{for } j = 1, 2, \ldots n-1, \tag{A.9a}$$

$$1 + \alpha_n R_f \qquad \text{paid at } t_n. \tag{A.9b}$$

Here the coefficient $\alpha_j = \alpha(t_{j-1}, t_j)$ is the year fraction between $t_{j-1}$ and $t_j$. Year fractions for semiannual legs are often not exactly $\frac{1}{2}$ due to the vagaries of holidays, weekends, and shorter or longer months. They are determined by the swap's *day count basis*, which is typically 30/360 for USD legs, actual/365 for Sterling legs, and 30/360E for Euro legs. The final payment includes repayment of the notional, which is assumed to be 1.

On any date $t$, the value of the fixed leg is clearly the same as a collection of zero coupon bonds:

$$V_{fix}(t) = R_f \sum_{j=1}^{n} \alpha_j Z(t; t_j) + Z(t_n). \tag{A.10}$$

The actual dates $s_1, s_2, \ldots s_m = t_n$ of the floating leg are set up in exactly the same way, save that floating legs are usually quarterly instead of semiannual. The floating rate pays

$$\beta_j R_j^{fl} \qquad \text{paid at } s_j \qquad \text{for } j = 1, 2, \ldots m-1, \tag{A.11a}$$

$$1 + \beta_n R_m^{fl} \qquad \text{paid at } s_m, \tag{A.11b}$$

where $\beta_j$ is the day count fraction of the $j^{th}$ interval according to the floating leg's day count basis. The interest rate $R_j^{fl}$ for the $j^{th}$ interval $s_{j-1} < t \leq s_j$ is set *spot-lag* (generally 2) business days before the interval starts at $s_{j-1}$.

Theoretically, the value of a floating leg is the same as the value of 1 paid at the start date $t_0$ of the floating leg:

$$V_{fl}(t) = Z(t; t_0). \tag{A.12}$$

This, because if one starts with a single unit at $t_0$, one can invest it at the rate $R_1^{fl}$ from $t_0$ until $s_1$, one can then reinvest 1 at the new rate $R_2^{fl}$ from $s_1$ to $s_2$, …, and finally reinvest it at $R_m^{fl}$ for the final interval $s_{m-1}$ to $s_m$. Thus, starting with 1 unit of currency enables one to reconstruct the floating leg payments exactly.

A (payer) swap exchanges a fixed leg for a floating leg. Its value is

$$V_{sw}(t) = Z(t; t_0) - Z(t; t_n) - R_f \sum_{j=1}^{n} \alpha_j Z(t; t_j). \tag{A.13}$$

The swap rate is defined to be the value of $R_f$ which sets the value of this swap to 0. So at any date $t$, the swap rate is

$$R_s(t) = \frac{Z(t_0) - Z(t_n)}{L(t)}, \tag{A.14a}$$

where

$$L(t)=\sum_{j=1}^{n}\alpha_j Z(t;t_j) \tag{A.14b}$$

is the *level* (also called the *annuity* or *PV01*). The value of the swap can now be written in terms of the swap rate as

$$V_{sw}(t)=[R_s(t)-R_f]\,L(t). \tag{A.14c}$$

Consider a swaption, a European option on the swap. To price the option, we note that the level is clearly a tradeable instrument, since it is just a collection of zero coupon bonds, so we choose it as our numeraire. The fundamental theorem ensures that there is a probability measure $Q$, known as the level measure, such that for any tradeable instrument, the instrument's value $V(t)$ divided by $L(t)$ is a Martingale. In particular, the value of the swaption is

$$\frac{V_{swptn}(t)}{L(t)}=E^Q\left\{\frac{V_{call}(T)}{L(T)}|\mathfrak{F}_t\right\}\qquad \text{for all } T>t. \tag{A.15}$$

The simplest date $T$ for evaluating this expression is the swaption's exercise date. On the exercise date, the value of the swaption is clearly $[V_{sw}(t_{ex})]^+$, which is $[R_s(t_{ex})-R_f]^+\,L(t_{ex})$. Therefore the swaption value can be written as

$$V_{swptn}(t)=L(t)E^Q\left\{[R_s(t_{ex})-R_f]^+\,|\mathfrak{F}_t\right\} \tag{A.16}$$

where the expectation is over the level measure. Moreover, since $Z(t_0)-Z(t_n)$ represents a tradeable security (long a zero coupon bond, and short a second zero coupon bond), and $L(t)$ is our numeraire, the swap rate $R_s(t)$ is a Martingale in this measure. So from the Martingale representation theorem, we again conclude that

$$dR_s=C(t,*)dW, \tag{A.17}$$

for some coefficient $C\,(t,*)$, where $dW$ is Brownian motion.

Finally, letting $t=0$ be today, and and noting that today's zero coupon bond values $Z(0;t_j)$ are just the discount factors $D(t_j)$, we can write today's swaption value as

$$V_{swptn}(0)=L_0E^Q\left\{[R_s(t_{ex})-R_f]^+\,|\mathfrak{F}_0\right\}, \tag{A.18a}$$

where the swap rate $R_s(t)$ is a Martingale,

$$dR_s=C(t,*)dW, \tag{A.18b}$$

and today's level $L_0$ and today's swap rate are given by

$$L_0=\sum_{j=1}^{n}\alpha_j D(t_j),\qquad R_s(0)=\frac{D(t_0)-D(t_n)}{L_0}. \tag{A.18c}$$

This is clearly the same theoretical framework as the general European option above.

**Appendix B. Analysis of the SABR model.**

Here we use singular perturbation techniques to price European options under the SABR model. Our analysis is based on a small volatility expansion, where we take both the volatility $\hat{\alpha}$ and the "volvol" $\nu$ to be small. To carry out this analysis in a systematic fashion, we re-write $\hat{\alpha} \longrightarrow \varepsilon\hat{\alpha}$, and $\nu \longrightarrow \varepsilon\nu$, and analyze

$$d\hat{F} = \varepsilon\hat{\alpha}C(\hat{F})dW_1, \tag{B.1a}$$

$$d\hat{\alpha} = \varepsilon\nu\hat{\alpha}dW_2, \tag{B.1b}$$

with

$$dW_1 dW_2 = \rho dt, \tag{B.1c}$$

in the limit $\varepsilon \ll 1$. This is the *distinguished limit* [22], [23] in the language of singular perturbation theory. After obtaining the results we replace $\varepsilon\hat{\alpha} \longrightarrow \hat{\alpha}$, and $\varepsilon\nu \longrightarrow \nu$ to get the answer in terms of the original variables. We first analyze the model with a general $C(\hat{F})$, and then specialize the results to the power law $\hat{F}^\beta$. This is notationally simpler than working with the power law throughout, and the more general result may prove valuable in some future application.

We first use the forward Kolmogorov equation to simplify the option pricing problem. Suppose the economy is in state $\hat{F}(t) = f,\ \hat{\alpha}(t) = \alpha$ at date $t$. Define the probability density $p(t, f, \alpha; T, F, A)$ by

$$p(t, f, \alpha; T, F, A)dFdA = \text{prob}\left\{ F < \hat{F}(T) < F + dF,\ A < \hat{\alpha}(T) < A + dA \,\middle|\, \hat{F}(t) = f,\ \hat{\alpha}(t) = \alpha \right\}. \tag{B.2}$$

As a function of the *forward variables* $T, F, A$, the density $p$ satisfies the forward Kolmogorov equation (the Főkker-Planck equation)

$$p_T = \tfrac{1}{2}\varepsilon^2 A^2 [C^2(F)p]_{FF} + \varepsilon^2\rho\nu[A^2 C(F)p]_{FA} + \tfrac{1}{2}\varepsilon^2\nu^2[A^2 p]_{AA} \qquad \text{for } T > t, \tag{B.3a}$$

with

$$p = \delta(F - f)\delta(A - \alpha) \qquad \text{at } T = t, \tag{B.3b}$$

as is well-known [25], [26], [27]. Here, and throughout, we use subscripts to denote partial derivatives.

Let $V(t, f, \alpha)$ be the value of a European call option at date $t$, when the economy is in state $\hat{F}(t) = f,\ \hat{\alpha}(t) = \alpha$. Let $t_{ex}$ be the option's exercise date, and let $K$ be its strike. Omitting the discount factor $D(t_{set})$, which factors out exactly, the value of the option is

$$\begin{aligned} V(t, f, \alpha) &= E\left\{ [\hat{F}(t_{ex}) - K]^+ \mid \hat{F}(t) = f,\ \hat{\alpha}(t) = \alpha \right\} \\ &= \int_{-\infty}^{\infty}\int_K^{\infty} (F - K)p(t, f, \alpha; t_{ex}, F, A)dFdA. \end{aligned} \tag{B.4}$$

See 2.1a. Since

$$p(t, f, \alpha; t_{ex}, F, A) = \delta(F - f)\delta(A - \alpha) + \int_t^{t_{ex}} p_T(t, f, \alpha; T, F, A)dT, \tag{B.5}$$

we can re-write $V(t, f, \alpha)$ as

$$V(t, f, \alpha) = [f - K]^+ + \int_t^{t_{ex}}\int_K^{\infty}\int_{-\infty}^{\infty} (F - K)p_T(t, f, \alpha; T, F, A)dAdFdT. \tag{B.6}$$

We substitute B.3a for $p_T$ into B.6. Integrating the $A$ derivatives $\varepsilon^2\rho\nu[A^2C(F)p]_{FA}$ and $\frac{1}{2}\varepsilon^2\nu^2[A^2p]_{AA}$ over all $A$ yields zero. Therefore our option price reduces to

$$V(t,f,\alpha) = [f-K]^+ + \tfrac{1}{2}\varepsilon^2 \int_t^{t_{ex}} \int_{-\infty}^{\infty} \int_K^{\infty} A^2\,(F-K)\,[C^2(F)p]_{FF}\,dFdAdT, \tag{B.7}$$

where we have switched the order of integration. Integrating by parts twice with respect to $F$ now yields

$$V(t,f,\alpha) = [f-K]^+ + \tfrac{1}{2}\varepsilon^2 C^2(K) \int_t^{t_{ex}} \int_{-\infty}^{\infty} A^2 p(t,f,\alpha;T,K,A)\,dAdT. \tag{B.8}$$

The problem can be simplified further by defining

$$P(t,f,\alpha;T,K) = \int_{-\infty}^{\infty} A^2 p(t,f,\alpha;T,K,A)\,dA. \tag{B.9}$$

Then $P$ satisfies the backward's Kolmogorov equation [25], [26], [27]

$$P_t + \tfrac{1}{2}\varepsilon^2\alpha^2C^2(f)P_{ff} + \varepsilon^2\rho\nu\alpha^2C(f)P_{f\alpha} + \tfrac{1}{2}\varepsilon^2\nu^2\alpha^2P_{\alpha\alpha} = 0, \qquad \text{for } t<T \tag{B.10a}$$

$$P = \alpha^2\delta(f-K), \qquad \text{for } t=T. \tag{B.10b}$$

Since $t$ does not appear explicitly in this equation, $P$ depends only on the combination $T-t$, and not on $t$ and $T$ separately. So define

$$\tau = T-t, \qquad \tau_{ex} = t_{ex}-t. \tag{B.11}$$

Then our pricing formula becomes

$$V(t,f,\alpha) = [f-K]^+ + \tfrac{1}{2}\varepsilon^2 C^2(K) \int_0^{\tau_{ex}} P(\tau,f,\alpha;K)\,d\tau \tag{B.12}$$

where $P(\tau,f,\alpha;K)$ is the solution of the problem

$$P_\tau = \tfrac{1}{2}\varepsilon^2\alpha^2C^2(f)P_{ff} + \varepsilon^2\rho\nu\alpha^2C(f)P_{f\alpha} + \tfrac{1}{2}\varepsilon^2\nu^2\alpha^2P_{\alpha\alpha}, \qquad \text{for } \tau>0, \tag{B.13a}$$

$$P = \alpha^2\delta(f-K), \qquad \text{for } \tau=0. \tag{B.13b}$$

In this appendix we solve B.13a, B.13b to obtain $P(\tau,f,\alpha;K)$, and then substitute this solution into B.12 to obtain the option value $V(t,f,\alpha)$. This yields the option price under the SABR model, but the resulting formulas are awkward and not very useful. To cast the results in a more usable form, we re-compute the option price under the normal model

$$d\hat{F} = \sigma_N dW, \tag{B.14a}$$

and then equate the two prices to determine which normal volatility $\sigma_N$ needs to be used to reproduce the option's price under the SABR model. That is, we find the "implied normal volatility" of the option under the SABR model. By doing a second comparison between option prices under the log normal model

$$d\hat{F} = \sigma_B \hat{F} dW \tag{B.14b}$$

and the normal model, we then convert the implied normal volatility to the usual implied log-normal (Black-Scholes) volatility. That is, we quote the option price predicted by the SABR model in terms of the option's implied volatility.

**B.1. Singular perturbation expansion.** Using a straightforward perturbation expansion would yield a Gaussian density to leading order,

$$P = \frac{\alpha}{\sqrt{2\pi\varepsilon^2 C^2(K)\tau}} e^{-\frac{(f-K)^2}{2\varepsilon^2\alpha^2 C^2(K)\tau}} \{1 + \cdots\}. \tag{B.15a}$$

Since the "$+\cdots$" involves powers of $(f-K)/\varepsilon\alpha C(K)$, this expansion would become inaccurate as soon as $(f-K)C'(K)/C(K)$ becomes a significant fraction of 1; i.e., as soon as $C(f)$ and $C(K)$ are significantly different. Stated differently, small changes in the exponent cause much greater changes in the probability density. A better approach is to re-cast the series as

$$P = \frac{\alpha}{\sqrt{2\pi\varepsilon^2 C^2(K)\tau}} e^{-\frac{(f-K)^2}{2\varepsilon^2\alpha^2 C^2(K)\tau}\{1+\cdots\}} \tag{B.15b}$$

and expand the exponent, since one expects that only small changes to the exponent will be needed to effect the much larger changes in the density. This expansion also describes the basic physics better — $P$ is essentially a Gaussian probability density which tails off faster or slower depending on whether the "diffusion coefficient" $C(f)$ decreases or increases.

We can refine this approach by noting that the exponent is the integral

$$\frac{(f-K)^2}{2\varepsilon^2\alpha^2 C^2(K)\tau}\{1+\cdots\} = \frac{1}{2\tau}\left(\frac{1}{\varepsilon\alpha}\int_K^f \frac{df'}{C(f')}\right)^2 \{1+\cdots\}. \tag{B.16}$$

Suppose we define the new variable

$$z = \frac{1}{\varepsilon\alpha}\int_K^f \frac{df'}{C(f')}. \tag{B.17}$$

so that the solution $P$ is essentially $e^{-z^2/2}$. To leading order, the density is Gaussian in the variable $z$, which is determined by how "easy" or "hard" it is to diffuse from $K$ to $f$, which closely matches the underlying physics. The fact that the Gaussian changes by orders of magnitude as $z^2$ increases should be largely irrelevent to the quality of the expansion. This approach is directly related to the geometric optics technique that is so successful in wave propagation and quantum electronics [28], [23]. To be more specific, we shall use the near identity transform method to carry out the geometric optics expansion. This method, pioneered in [29], transforms the problem order-by-order into a simple canonical problem, which can then be solved trivially. Here we obtain the solution only through $O(\varepsilon^2)$, truncating all higher order terms.

Let us change variables from $f$ to

$$z = \frac{1}{\varepsilon\alpha}\int_K^f \frac{df'}{C(f')}, \tag{B.18a}$$

and to avoid confusion, we define

$$B(\varepsilon\alpha z) = C(f). \tag{B.18b}$$

Then

$$\frac{\partial}{\partial f} \longrightarrow \frac{1}{\varepsilon\alpha C(f)}\frac{\partial}{\partial z} = \frac{1}{\varepsilon\alpha B(\varepsilon\alpha z)}\frac{\partial}{\partial z}, \qquad \frac{\partial}{\partial\alpha} \longrightarrow \frac{\partial}{\partial\alpha} - \frac{z}{\alpha}\frac{\partial}{\partial z}, \tag{B.19a}$$

and

$$\frac{\partial^2}{\partial f^2} \longrightarrow \frac{1}{\varepsilon^2\alpha^2 B^2(\varepsilon\alpha z)}\left\{\frac{\partial^2}{\partial z^2} - \varepsilon\alpha\frac{B'(\varepsilon\alpha z)}{B(\varepsilon\alpha z)}\frac{\partial}{\partial z}\right\}, \tag{B.19b}$$

$$\frac{\partial^2}{\partial f\partial\alpha} \longrightarrow \frac{1}{\varepsilon\alpha B(\varepsilon\alpha z)}\left\{\frac{\partial^2}{\partial z\partial\alpha} - \frac{z}{\alpha}\frac{\partial^2}{\partial z^2} - \frac{1}{\alpha}\frac{\partial}{\partial z}\right\}, \tag{B.19c}$$

$$\frac{\partial^2}{\partial\alpha^2} \longrightarrow \frac{\partial^2}{\partial\alpha^2} - \frac{2z}{\alpha}\frac{\partial^2}{\partial z\partial\alpha} + \frac{z^2}{\alpha^2}\frac{\partial^2}{\partial z^2} + \frac{2z}{\alpha^2}\frac{\partial}{\partial z}. \tag{B.19d}$$

Also,

$$\delta(f-K) = \delta(\varepsilon\alpha z C(K)) = \frac{1}{\varepsilon\alpha C(K)}\delta(z). \tag{B.19e}$$

Therefore, B.12 through B.13b become

$$V(t,f,a) = [f-K]^+ + \tfrac{1}{2}\varepsilon^2 C^2(K)\int_0^{\tau_{ex}} P(\tau,z,\alpha)\,d\tau, \tag{B.20}$$

where $P(\tau,z,\alpha)$ is the solution of

$$\begin{aligned} P_\tau = \tfrac{1}{2}\left(1-2\varepsilon\rho\nu z+\varepsilon^2\nu^2 z^2\right)P_{zz} - \tfrac{1}{2}\varepsilon a\frac{B'}{B}P_z + (\varepsilon\rho\nu-\varepsilon^2\nu^2 z)(\alpha P_{z\alpha}-P_z) \\ +\tfrac{1}{2}\varepsilon^2\nu^2\alpha^2 P_{\alpha a} \qquad \text{for } \tau>0 \end{aligned} \tag{B.21a}$$

$$P = \frac{\alpha}{\varepsilon C(K)}\delta(z) \qquad \text{at } \tau=0. \tag{B.21b}$$

Accordingly, let us define $\hat{P}(\tau,z,\alpha)$ by

$$\hat{P} = \frac{\varepsilon}{\alpha}C(K)P. \tag{B.22}$$

In terms of $\hat{P}$, we obtain

$$V(t,f,a) = [f-K]^+ + \tfrac{1}{2}\varepsilon\alpha C(K)\int_0^{\tau_{ex}} \hat{P}(\tau,z,\alpha)\,d\tau, \tag{B.23}$$

where $\hat{P}(\tau,z,\alpha)$ is the solution of

$$\begin{aligned} \hat{P}_\tau = \tfrac{1}{2}\left(1-2\varepsilon\rho\nu z+\varepsilon^2\nu^2 z^2\right)\hat{P}_{zz} - \tfrac{1}{2}\varepsilon a\frac{B'}{B}\hat{P}_z + (\varepsilon\rho\nu-\varepsilon^2\nu^2 z)\alpha\hat{P}_{z\alpha} \\ +\tfrac{1}{2}\varepsilon^2\nu^2(\alpha^2\hat{P}_{\alpha\alpha}+2\alpha\hat{P}_\alpha) \qquad \text{for } \tau>0, \end{aligned} \tag{B.24a}$$

$$\hat{P} = \delta(z) \qquad \text{at } \tau=0. \tag{B.24b}$$

To leading order $\hat{P}$ is the solution of the standard diffusion problem $\hat{P}_\tau = \frac{1}{2}\hat{P}_{zz}$ with $\hat{P}=\delta(z)$ at $\tau=0$. So it is a Gaussian to leading order. The next stage is to transform the problem to the standard diffusion problem through $O(\varepsilon)$, and then through $O(\varepsilon^2)$, .... This is the near identify transform method which has proven so powerful in near-Hamiltonian systems [29].

Note that the variable $\alpha$ does not enter the problem for $\hat{P}$ until $O(\varepsilon)$, so

$$\hat{P}(\tau, z, \alpha) = \hat{P}_0(\tau, z) + \hat{P}_1(\tau, z, \alpha) + \cdots \tag{B.25}$$

Consequently, the derivatives $\hat{P}_{z\alpha}$, $\hat{P}_{\alpha\alpha}$, and $\hat{P}_\alpha$ are all $O(\varepsilon)$. Recall that we are only solving for $\hat{P}$ through $O(\varepsilon^2)$. So, through this order, we can re-write our problem as

$$\hat{P}_\tau = \tfrac{1}{2}\left(1 - 2\varepsilon\rho\nu z + \varepsilon^2\nu^2 z^2\right)\hat{P}_{zz} - \tfrac{1}{2}\varepsilon a \frac{B'}{B}\hat{P}_z + \varepsilon\rho\nu\alpha\hat{P}_{z\alpha} \qquad \text{for } \tau > 0 \tag{B.26a}$$

$$\hat{P} = \delta(z) \qquad \text{at } \tau = 0. \tag{B.26b}$$

Let us now eliminate the $\frac{1}{2}\varepsilon a(B'/B)\hat{P}_z$ term. Define $H(\tau, z, \alpha)$ by

$$\hat{P} = \sqrt{C(f)/C(K)}H \equiv \sqrt{B(\varepsilon\alpha z)/B(0)}H. \tag{B.27}$$

Then

$$\hat{P}_z = \sqrt{B(\varepsilon\alpha z)/B(0)}\left\{H_z + \tfrac{1}{2}\varepsilon\alpha\frac{B'}{B}H\right\}, \tag{B.28a}$$

$$\hat{P}_{zz} = \sqrt{B(\varepsilon\alpha z)/B(0)}\left\{H_{zz} + \varepsilon\alpha\frac{B'}{B}H_z + \varepsilon^2\alpha^2\left[\frac{B''}{2B} - \frac{B'^2}{4B^2}\right]H\right\}, \tag{B.28b}$$

$$\hat{P}_{z\alpha} = \sqrt{B(\varepsilon\alpha z)/B(0)}\left\{H_{z\alpha} + \tfrac{1}{2}\varepsilon z\frac{B'}{B}H_z + \tfrac{1}{2}\varepsilon\alpha\frac{B'}{B}H_\alpha + \tfrac{1}{2}\varepsilon\frac{B'}{B}H + O(\varepsilon^2)\right\}. \tag{B.28c}$$

The option price now becomes

$$V(t, f, a) = [f - K]^+ + \tfrac{1}{2}\varepsilon\alpha\sqrt{B(0)B(\varepsilon\alpha z)}\int_0^{\tau_{ex}} H(\tau, z, \alpha)\, d\tau, \tag{B.29}$$

where

$$\begin{aligned} H_\tau = {} & \tfrac{1}{2}\left(1 - 2\varepsilon\rho\nu z + \varepsilon^2\nu^2 z^2\right)H_{zz} - \tfrac{1}{2}\varepsilon^2\rho\nu\alpha\frac{B'}{B}(zH_z - H) \\ & + \varepsilon^2\alpha^2\left(\tfrac{1}{4}\frac{B''}{B} - \tfrac{3}{8}\frac{B'^2}{B^2}\right)H + \varepsilon\rho\nu\alpha(H_{z\alpha} + \tfrac{1}{2}\varepsilon\alpha\frac{B'}{B}H_\alpha) \qquad \text{for } \tau > 0 \end{aligned} \tag{B.30a}$$

$$H = \delta(z) \qquad \text{at } \tau = 0. \tag{B.30b}$$

Equations B.30a, B.30b are independent of $\alpha$ to leading order, and at $O(\varepsilon)$ they depend on $\alpha$ only through the last term $\varepsilon\rho\nu\alpha(H_{z\alpha} + \frac{1}{2}\varepsilon\alpha\frac{B'}{B}H_\alpha)$. As above, since B.30a is independent of $\alpha$ to leading order, we can conclude that the $\alpha$ derivatives $H_\alpha$ and $H_{z\alpha}$ are no larger than $O(\varepsilon)$, and so the last term is actually no larger than $O(\varepsilon^2)$. Therefore $H$ is independent of $\alpha$ until $O(\varepsilon^2)$ and the $\alpha$ derivatives are actually no larger than $O(\varepsilon^2)$. Thus, the last term is actually only $O(\varepsilon^3)$, and can be neglected since we are only working through $O(\varepsilon^2)$. So,

$$H_\tau = \tfrac{1}{2}\left(1 - 2\varepsilon\rho\nu z + \varepsilon^2\nu^2 z^2\right)H_{zz} - \tfrac{1}{2}\varepsilon^2\rho\nu\alpha\frac{B'}{B}(zH_z - H) + \varepsilon^2\alpha^2\left(\tfrac{1}{4}\frac{B''}{B} - \tfrac{3}{8}\frac{B'^2}{B^2}\right)H \qquad \text{for } \tau > 0 \tag{B.31a}$$

$$H = \delta(z) \qquad \text{at } \tau = 0. \tag{B.31b}$$

There are no longer any $\alpha$ derivatives, so we can now treat $\alpha$ as a parameter instead of as an independent variable. That is, we have succeeded in effectively reducing the problem to one dimension.

Let us now remove the $H_z$ term through $O(\varepsilon^2)$. To leading order, $B'(\varepsilon\alpha z)/B(\varepsilon\alpha z)$ and $B''(\varepsilon\alpha z)/B(\varepsilon\alpha z)$ are constant. We can replace these ratios by

$$b_1 = B'(\varepsilon\alpha z_0)/B(\varepsilon\alpha z_0), \qquad b_2 = B''(\varepsilon\alpha z_0)/B(\varepsilon\alpha z_0), \tag{B.32}$$

commiting only an $O(\varepsilon)$ error, where the constant $z_0$ will be chosen later. We now define $\hat{H}$ by

$$H = e^{\varepsilon^2\rho\nu\alpha b_1 z^2/4}\hat{H}. \tag{B.33}$$

Then our option price becomes

$$V(t,f,a) = [f-K]^+ + \tfrac{1}{2}\varepsilon\alpha\sqrt{B(0)B(\varepsilon\alpha z)}e^{\varepsilon^2\rho\nu\alpha b_1 z^2/4}\int_0^{\tau_{ex}} \hat{H}(\tau,z)\,d\tau, \tag{B.34}$$

where $\hat{H}$ is the solution of

$$\hat{H}_\tau = \tfrac{1}{2}\left(1 - 2\varepsilon\rho\nu z + \varepsilon^2\nu^2 z^2\right)\hat{H}_{zz} + \varepsilon^2\alpha^2\left(\tfrac{1}{4}b_2 - \tfrac{3}{8}b_1^2\right)\hat{H} + \tfrac{3}{4}\varepsilon^2\rho\nu\alpha b_1\hat{H} \qquad \text{for } \tau > 0 \tag{B.35a}$$

$$\hat{H} = \delta(z) \qquad \text{at } \tau = 0. \tag{B.35b}$$

We've almost beaten the equation into shape. We now define

$$x = \frac{1}{\varepsilon\nu}\int_0^{\varepsilon\nu z}\frac{d\zeta}{\sqrt{1-2\rho\zeta+\zeta^2}} = \frac{1}{\varepsilon\nu}\log\left(\frac{\sqrt{1-2\varepsilon\rho\nu z+\varepsilon^2\nu^2 z^2}-\rho+\varepsilon\nu z}{1-\rho}\right), \tag{B.36a}$$

which can be written implicitly as

$$\varepsilon\nu z = \sinh\varepsilon\nu x - \rho(\cosh\varepsilon\nu x - 1). \tag{B.36b}$$

In terms of x, our problem is

$$V(t,f,a) = [f-K]^+ + \tfrac{1}{2}\varepsilon\alpha\sqrt{B(0)B(\varepsilon\alpha z)}e^{\varepsilon^2\rho\nu\alpha b_1 z^2/4}\int_0^{\tau_{ex}} \hat{H}(\tau,x)\,d\tau, \tag{B.37}$$

with

$$\hat{H}_\tau = \tfrac{1}{2}\hat{H}_{xx} - \tfrac{1}{2}\varepsilon\nu I'(\varepsilon\nu z)\hat{H}_x + \varepsilon^2\alpha^2\left(\tfrac{1}{4}b_2 - \tfrac{3}{8}b_1^2\right)\hat{H} + \tfrac{3}{4}\varepsilon^2\rho\nu\alpha b_1\hat{H} \qquad \text{for } \tau > 0 \tag{B.38a}$$

$$\hat{H} = \delta(x) \qquad \text{at } \tau = 0. \tag{B.38b}$$

Here

$$I(\zeta) = \sqrt{1-2\rho\zeta+\zeta^2}. \tag{B.39}$$

The final step is to define $Q$ by

$$\hat{H} = I^{1/2}(\varepsilon\nu z(x))Q = \left(1-2\varepsilon\rho\nu z+\varepsilon^2\nu^2 z^2\right)^{1/4}Q. \tag{B.40}$$

Then

$$\hat{H}_x = I^{1/2}(\varepsilon\nu z)\left[Q_x + \tfrac{1}{2}\varepsilon\nu I'(\varepsilon\nu z)Q\right], \tag{B.41a}$$

$$\hat{H}_{xx} = I^{1/2}(\varepsilon\nu z)\left[Q_{xx} + \varepsilon\nu I' Q_x + \varepsilon^2\nu^2\left(\tfrac{1}{2}I''I + \tfrac{1}{4}I'I'\right)Q\right], \tag{B.41b}$$

and so

$$V(t,f,a) = [f-K]^+ + \tfrac{1}{2}\varepsilon\alpha\sqrt{B(0)B(\varepsilon\alpha z)}I^{1/2}(\varepsilon\nu z)e^{\frac{1}{4}\varepsilon^2\rho\nu\alpha b_1 z^2}\int_0^{\tau_{ex}} Q(\tau,x)\,d\tau, \tag{B.42}$$

where $Q$ is the solution of

$$Q_\tau = \tfrac{1}{2}Q_{xx} + \varepsilon^2\nu^2\left(\tfrac{1}{4}I''I - \tfrac{1}{8}I'I'\right)Q + \varepsilon^2\alpha^2\left(\tfrac{1}{4}b_2 - \tfrac{3}{8}b_1^2\right)Q + \tfrac{3}{4}\varepsilon^2\rho\nu\alpha b_1 Q \tag{B.43a}$$

for $\tau > 0$, with

$$Q = \delta(x) \qquad \text{at } \tau = 0. \tag{B.43b}$$

As above, we can replace $I(\varepsilon\nu z), I'(\varepsilon\nu z), I''(\varepsilon\nu z)$ by the constants $I(\varepsilon\nu z_0), I'(\varepsilon\nu z_0), I''(\varepsilon\nu z_0)$, and commit only $O(\varepsilon)$ errors. Define the constant $\kappa$ by

$$\kappa = \nu^2\left(\tfrac{1}{4}I''(\varepsilon\nu z_0)I(\varepsilon\nu z_0) - \tfrac{1}{8}\left[I'(\varepsilon\nu z_0\right]^2)\right) + \alpha^2\left(\tfrac{1}{4}b_2 - \tfrac{3}{8}b_1^2\right) + \tfrac{3}{4}\rho\nu\alpha b_1, \tag{B.44}$$

where $z_0$ will be chosen later. Then through $O(\varepsilon^2)$, we can simplify our equation to

$$Q_\tau = \tfrac{1}{2}Q_{xx} + \varepsilon^2\kappa Q \qquad \text{for } \tau > 0, \tag{B.45a}$$

$$Q = \delta(x) \qquad \text{at } \tau = 0. \tag{B.45b}$$

The solution of B.45a, B.45b is clearly

$$Q = \frac{1}{\sqrt{2\pi\tau}}e^{-x^2/2\tau}e^{\varepsilon^2\kappa\tau} = \frac{1}{\sqrt{2\pi\tau}}e^{-x^2/2\tau}\frac{1}{\left(1 - \frac{2}{3}\kappa\varepsilon^2\tau + \cdots\right)^{3/2}} \tag{B.46}$$

through $O(\varepsilon^2)$.

This solution yields the option price

$$V(t,f,a) = [f-K]^+ + \tfrac{1}{2}\varepsilon\alpha\sqrt{B(0)B(\varepsilon\alpha z)}I^{1/2}(\varepsilon\nu z)e^{\frac{1}{4}\varepsilon^2\rho\nu\alpha b_1 z^2}\int_0^{\tau_{ex}} \frac{1}{\sqrt{2\pi\tau}}e^{-x^2/2\tau}e^{\varepsilon^2\kappa\tau}\,d\tau. \tag{B.47}$$

Observe that this can be written as

$$V(t,f,a) = [f-K]^+ + \tfrac{1}{2}\frac{f-K}{x}\int_0^{\tau_{ex}} \frac{1}{\sqrt{2\pi\tau}}\,e^{-x^2/2\tau}e^{\varepsilon^2\theta}e^{\varepsilon^2\kappa\tau}\,d\tau, \tag{B.48a}$$

where

$$\varepsilon^2\theta = \log\left(\frac{\varepsilon\alpha z}{f-K}\sqrt{B(0)B(\varepsilon\alpha z)}\right) + \log\left(\frac{xI^{1/2}(\varepsilon\nu z)}{z}\right) + \tfrac{1}{4}\varepsilon^2\rho\nu\alpha b_1 z^2 \tag{B.48b}$$

Moreover, quite amazingly,

$$e^{\varepsilon^2\kappa\tau} = \frac{1}{\left(1 - \frac{2}{3}\kappa\varepsilon^2\tau\right)^{3/2}} = \frac{1}{\left(1 - 2\varepsilon^2\tau\dfrac{\theta}{x^2}\right)^{3/2}} + O(\varepsilon^4), \tag{B.48c}$$

through $O(\varepsilon^2)$. This can be shown by expanding $\varepsilon^2\theta$ through $O(\varepsilon^2)$, and noting that $\varepsilon^2\theta/x^2$matches $\kappa/3$. Therefore our option price is

$$V(t,f,a) = [f-K]^+ + \tfrac{1}{2}\frac{f-K}{x}\int_0^{\tau_{ex}} \frac{1}{\sqrt{2\pi\tau}}\, e^{-x^2/2\tau} e^{\varepsilon^2\theta} \frac{d\tau}{\left(1-\dfrac{2\tau}{x^2}\varepsilon^2\theta\right)^{3/2}}, \tag{B.49}$$

and changing integration variables to

$$q = \frac{x^2}{2\tau}, \tag{B.50}$$

reduces this to

$$V(t,f,a) = [f-K]^+ + \frac{|f-K|}{4\sqrt{\pi}}\int_{\frac{x^2}{2\tau_{ex}}}^{\infty} \frac{e^{-q+\varepsilon^2\theta}}{(q-\varepsilon^2\theta)^{3/2}}dq. \tag{B.51}$$

That is, the value of a European call option is given by

$$V(t,f,a) = [f-K]^+ + \frac{|f-K|}{4\sqrt{\pi}}\int_{\frac{x^2}{2\tau_{ex}}-\varepsilon^2\theta}^{\infty} \frac{e^{-q}}{q^{3/2}}dq, \tag{B.52a}$$

with

$$\varepsilon^2\theta = \log\left(\frac{\varepsilon\alpha z}{f-K}\sqrt{B(0)B(\varepsilon\alpha z)}\right) + \log\left(\frac{xI^{1/2}(\varepsilon\nu z)}{z}\right) + \tfrac{1}{4}\varepsilon^2\rho\nu\alpha b_1 z^2, \tag{B.52b}$$

through $O(\varepsilon^2)$.

**B.2. Equivalent normal volatility.** Equations B.52a and B.52a are a formula for the dollar price of the call option under the SABR model. The utility and beauty of this formula is not overwhelmingly apparent. To obtain a useful formula, we convert this dollar price into the equivalent implied volatilities. We first obtain the implied *normal* volatility, and then the standard *log normal* (Black) volatility.

Suppose we repeated the above analysis for the ordinary *normal model*

$$d\hat{F} = \sigma_N dW, \qquad \hat{F}(0) = f. \tag{B.53a}$$

where the normal volatily $\sigma_N$ is constant, not stochastic. (This model is also called the *absolute* or *Gaussian* model). We would find that the option value for the normal model is exactly

$$V(t,f) = [f-K]^+ + \frac{|f-K|}{4\sqrt{\pi}}\int_{\frac{(f-K)^2}{2\sigma_N^2\tau_{ex}}}^{\infty} \frac{e^{-q}}{q^{3/2}}dq \tag{B.53b}$$

This can be seen by setting $C(f)$ to 1, setting $\varepsilon\alpha$ to $\sigma_N$ and setting $\nu$ to 0 in B.52a, B.52b. Working out this integral then yields the exact European option price

$$V(t,f) = (f-K)\mathcal{N}(\frac{f-K}{\sigma_N\sqrt{\tau_{ex}}}) + \sigma_N\sqrt{\tau_{ex}}\mathcal{G}(\frac{f-K}{\sigma_N\sqrt{\tau_{ex}}}), \tag{B.54a}$$

for the normal model, where $\mathcal{N}$ is the normal distribution and $\mathcal{G}$ is the Gaussian density

$$\mathcal{G}(q) = \frac{1}{\sqrt{2\pi}}e^{-q^2/2}. \tag{B.54b}$$

From B.53b it is clear that the option price under the normal model matches the option price under the SABR model B.52a, B.52a if and only if we choose the normal volatility $\sigma_N$ to be

$$\frac{1}{\sigma_N^2} = \frac{x^2}{(f-K)^2}\left\{1 - 2\varepsilon^2 \frac{\theta}{x^2}\tau_{ex}\right\}. \tag{B.55}$$

Taking the square root now shows the option's *implied normal* (absolute) *volatility* is given by

$$\sigma_N = \frac{f-K}{x}\left\{1 + \varepsilon^2 \frac{\theta}{x^2}\tau_{ex} + \cdots\right\} \tag{B.56}$$

through $O(\varepsilon^2)$.

Before continuing to the implied *log normal* volatility, let us seek the simplest possible way to re-write this answer which is correct through $O(\varepsilon^2)$. Since $x = z[1 + O(\varepsilon)]$, we can re-write the answer as

$$\sigma_N = \left(\frac{f-K}{z}\right)\left(\frac{z}{x(z)}\right)\left\{1 + \varepsilon^2\left(\phi_1 + \phi_2 + \phi_3\right)\tau_{ex} + \cdots\right\}, \tag{B.57a}$$

where

$$\frac{f-K}{z} = \frac{\varepsilon\alpha(f-K)}{\int_K^f \frac{df'}{C(f')}} = \left(\frac{1}{f-K}\int_K^f \frac{df'}{\varepsilon\alpha C(f')}\right)^{-1}.$$

This factor represents the average difficulty in diffusing from today's forward $f$ to the strike $K$, and would be present even if the volatility were not stochastic.

The next factor is

$$\frac{z}{x(z)} = \frac{\zeta}{\log\left(\frac{\sqrt{1-2\rho\zeta+\zeta^2} - \rho + \zeta}{1-\rho}\right)}, \tag{B.57b}$$

where

$$\zeta = \varepsilon\nu z = \frac{\nu}{\alpha}\int_K^f \frac{df'}{C(f')} = \frac{\nu}{\alpha}\frac{f-K}{C(f_{av})}\left\{1 + O(\varepsilon^2)\right\}. \tag{B.57c}$$

Here $f_{av} = \sqrt{fK}$ is the geometric average of $f$ and $K$. (The arithmetic average could have been used equally well at this order of accuracy). This factor represents the main effect of the stochastic volatility.

The coefficients $\phi_1$, $\phi_2$, and $\phi_3$ provide relatively minor corrections. Through $O(\varepsilon^2)$ these corrections are

$$\varepsilon^2\phi_1 = \frac{1}{z^2}\log\left(\frac{\varepsilon\alpha z}{f-K}\sqrt{C(f)C(K)}\right) = \frac{2\gamma_2 - \gamma_1^2}{24}\varepsilon^2\alpha^2 C^2\left(f_{av}\right) + \cdots \tag{B.57d}$$

$$\varepsilon^2\phi_2 = \frac{1}{z^2}\log\left(\frac{x}{z}\left[1 - 2\varepsilon\rho\nu z + \varepsilon^2\nu^2 z^2\right]^{1/4}\right) = \frac{2-3\rho^2}{24}\varepsilon^2\nu^2 + \cdots \tag{B.57e}$$

$$\varepsilon^2\phi_3 = \tfrac{1}{4}\varepsilon^2\rho\alpha\nu\frac{B'(\varepsilon\nu z_0)}{B(\varepsilon\nu z_0)} = \tfrac{1}{4}\varepsilon^2\rho\nu\alpha\gamma_1 C(f_{av}) + \cdots \tag{B.57f}$$

where

$$\gamma_1 = \frac{C'(f_{av})}{C(f_{av})}, \qquad \gamma_2 = \frac{C''(f_{av})}{C(f_{av})}. \tag{B.57g}$$

Let us briefly summarize before continuing. Under the *normal model*, the value of a European call option with strike $K$ and exercise date $\tau_{ex}$ is given by B.54a, B.54b. For the SABR model,

(B.58a) $$d\hat{F} = \varepsilon\hat{\alpha}C(\hat{F})dW_1, \qquad \hat{F}(0) = f$$

(B.58b) $$d\hat{\alpha} = \varepsilon\nu\hat{\alpha}dW_2, \qquad \hat{\alpha}(0) = \alpha$$

(B.58c) $$dW_1 dW_2 = \rho dt,$$

the value of the call option is given by the same formula, at least through $O(\varepsilon^2)$, provided we use the implied normal volatility

(B.59a) $$\sigma_N(K) = \frac{\varepsilon\alpha(f-K)}{\int_K^f \frac{df'}{C(f')}} \cdot \left(\frac{\zeta}{\hat{x}(\zeta)}\right) \cdot \left\{1 + \left[\frac{2\gamma_2 - \gamma_1^2}{24}\alpha^2 C^2\left(f_{av}\right) + \tfrac{1}{4}\rho\nu\alpha\gamma_1 C\left(f_{av}\right) + \frac{2-3\rho^2}{24}\nu^2\right]\varepsilon^2\tau_{ex} + \cdots\right\}.$$

Here

(B.59b) $$f_{av} = \sqrt{fK}, \qquad \gamma_1 = \frac{C'(f_{av})}{C(f_{av})}, \qquad \gamma_2 = \frac{C''(f_{av})}{C(f_{av})},$$

(B.59c) $$\zeta = \frac{\nu}{\alpha}\frac{f-K}{C(f_{av})}, \qquad \hat{x}(\zeta) = \log\left(\frac{\sqrt{1-2\rho\zeta+\zeta^2} - \rho + \zeta}{1-\rho}\right).$$

The first two factors provide the dominant behavior, with the remaining factor $1+[\cdots]\varepsilon^2\tau_{ex}$ usually providing corrections of around 1% or so.

One can repeat the analysis for a European put option, or simply use call/put parity. This shows that the value of the put option under the SABR model is

(B.60) $$V_{put}(f,\alpha,K) = (K-f)\mathcal{N}(\frac{K-f}{\sigma_N\sqrt{\tau_{ex}}}) + \sigma_N\sqrt{\tau_{ex}}\mathcal{G}(\frac{K-f}{\sigma_N\sqrt{\tau_{ex}}})$$

where the implied normal volatility $\sigma_N$ is given by the same formulas B.59a - B.59c as the call.

We can revert to the original units by replacing $\varepsilon\alpha \longrightarrow \alpha$, $\varepsilon\nu \longrightarrow \nu$ everywhere in the above formulas; this is equivalent to setting $\varepsilon$ to 1 everywhere.

**B.3. Equivalent Black vol.** With the exception of JPY traders, most traders prefer to quote prices in terms of Black (log normal) volatilities, rather than normal volatilities. To derive the implied Black volatility, consider Black's model

(B.61) $$d\hat{F} = \varepsilon\sigma_B\hat{F}dW, \qquad \hat{F}(0) = f,$$

where we have written the volatility as $\varepsilon\sigma_B$ to stay consistent with the preceding analysis. For Black's model, the value of a European call with strike $K$ and exercise date $\tau_{ex}$ is

(B.62a) $$V_{call} = f\mathcal{N}(d_1) - K\mathcal{N}(d_2),$$

(B.62b) $$V_{put} = V_{call} + D(t_{set})[K-f],$$

with

$$d_{1,2} = \frac{\log f/K \pm \frac{1}{2}\varepsilon^2\sigma_B^2\tau_{ex}}{\varepsilon\sigma_B\sqrt{\tau_{ex}}}, \tag{B.62c}$$

where we are omitting the overall factor $D(t_{set})$ as before.

We can obtain the implied *normal* volatility for Black's model by repeating the preceding analysis for the SABR model with $C(f) = f$ and $\nu = 0$. Setting $C(f) = f$ and $\nu = 0$ in B.59a -B.59c shows that the normal volatility is

$$\sigma_N(K) = \frac{\varepsilon\sigma_B(f-K)}{\log f/K}\left\{1 - \tfrac{1}{24}\varepsilon^2\sigma_B^2\tau_{ex} + \cdots\right\}. \tag{B.63}$$

through $O(\varepsilon^2)$. Indeed, in [15] it is shown that the implied normal volatility for Black's model is

$$\sigma_N(K) = \varepsilon\sigma_B\sqrt{fK}\,\frac{1 + \frac{1}{24}\log^2 f/K + \frac{1}{1920}\log^4 f/K + \cdots}{1 + \frac{1}{24}\left(1 - \frac{1}{120}\log^2 f/K\right)\varepsilon^2\sigma_B^2\tau_{ex} + \frac{1}{5760}\varepsilon^4\sigma_B^4\tau_{ex}^2 + \cdots} \tag{B.64}$$

through $O(\varepsilon^4)$. We can find the implied Black vol for the SABR model by setting $\sigma_N$ obtained from Black's model in equation B.63 equal to $\sigma_N$ obtained from the SABR model in B.59a - B.59c. Through $O(\varepsilon^2)$ this yields

$$\sigma_B(K) = \frac{\alpha\log f/K}{\int_K^f \frac{df'}{C(f')}}\cdot\left(\frac{\zeta}{\hat{x}(\zeta)}\right)\cdot \left\{1 + \left[\frac{2\gamma_2 - \gamma_1^2 + 1/f_{av}^2}{24}\alpha^2C^2(f_{av}) + \tfrac{1}{4}\rho\nu\alpha\gamma_1 C(f_{av}) + \frac{2-3\rho^2}{24}\nu^2\right]\varepsilon^2\tau_{ex} + \cdots\right\} \tag{B.65}$$

This is the main result of this article. As before, the implied log normal volatility for puts is the same as for calls, and this formula can be re-cast in terms of the original variables by simpley setting $\varepsilon$ to 1.

**B.4. Stochastic $\beta$ model.** As originally stated, the SABR model consists of the special case $C(f) = f^\beta$:

$$d\hat{F} = \varepsilon\hat{\alpha}\hat{F}^\beta dW_1, \qquad \hat{F}(0) = f \tag{B.66a}$$

$$d\hat{\alpha} = \varepsilon\nu\hat{\alpha}dW_2, \qquad \hat{\alpha}(0) = \alpha \tag{B.66b}$$

$$dW_1 dW_2 = \rho dt. \tag{B.66c}$$

Making this substitution in **??** - **??** shows that the implied *normal* volatility for this model is

$$\sigma_N(K) = \frac{\varepsilon\alpha(1-\beta)(f-K)}{f^{1-\beta} - K^{1-\beta}}\cdot\left(\frac{\zeta}{\hat{x}(\zeta)}\right)\cdot \left\{1 + \left[\frac{-\beta(2-\beta)\alpha^2}{24f_{av}^{2-2\beta}} + \frac{\rho\alpha\nu\beta}{4f_{av}^{1-\beta}} + \frac{2-3\rho^2}{24}\nu^2\right]\varepsilon^2\tau_{ex} + \cdots\right\} \tag{B.67a}$$

through $O(\varepsilon^2)$, where $f_{av} = \sqrt{fK}$ as before and

$$\zeta = \frac{\nu}{\alpha}\frac{f-K}{f_{av}^\beta}, \qquad \hat{x}(\zeta) = \log\left(\frac{\sqrt{1-2\rho\zeta+\zeta^2} - \rho + \zeta}{1-\rho}\right). \tag{B.67b}$$

We can simplify this formula by expanding

(B.68a) $$f - K = \sqrt{fK}\log f/K \left\{1 + \tfrac{1}{24}\log^2 f/K + \tfrac{1}{1920}\log^4 f/K + \cdots\right. ,$$

(B.68b) $$f^{1-\beta} - K^{1-\beta} = (1-\beta)(fK)^{(1-\beta)/2}\log f/K \left\{1 + \tfrac{(1-\beta)^2}{24}\log^2 f/K + \tfrac{(1-\beta)^4}{1920}\log^4 f/K + \cdots\right.$$

and neglecting terms higher than fourth order. This expansion reduces the implied normal volatility to

(B.69a) $$\sigma_N(K) = \varepsilon\alpha(fK)^{\beta/2}\frac{1 + \frac{1}{24}\log^2 f/K + \frac{1}{1920}\log^4 f/K + \cdots}{1 + \frac{(1-\beta)^2}{24}\log^2 f/K + \frac{(1-\beta)^4}{1920}\log^4 f/K + \cdots}\cdot\left(\frac{\zeta}{\hat{x}(\zeta)}\right)\cdot \left\{1 + \left[\frac{-\beta(2-\beta)\alpha^2}{24(fK)^{1-\beta}} + \frac{\rho\alpha\nu\beta}{4(fK)^{(1-\beta)/2}} + \frac{2-3\rho^2}{24}\nu^2\right]\varepsilon^2\tau_{ex} + \cdots\right\},$$

where

(B.69b) $$\zeta = \frac{\nu}{\alpha}(fK)^{(1-\beta)/2}\log f/K, \qquad \hat{x}(\zeta) = \log\left(\frac{\sqrt{1-2\rho\zeta+\zeta^2} - \rho + \zeta}{1-\rho}\right).$$

This is the formula we use in pricing European calls and puts.

To obtain the implied Black volatility, we equate the implied normal volatility $\sigma_N(K)$ for the SABR model obtained in B.69a - B.69b to the implied normal volatility for Black's model obtained in B.63. This shows that the implied Black volatility for the SABR model is

(B.69c) $$\sigma_B(K) = \frac{\varepsilon\alpha}{(fK)^{(1-\beta)/2}}\frac{1}{1 + \frac{(1-\beta)^2}{24}\log^2 f/K + \frac{(1-\beta)^4}{1920}\log^4 f/K + \cdots}\cdot\left(\frac{\zeta}{\hat{x}(\zeta)}\right)\cdot \left\{1 + \left[\frac{(1-\beta)^2\alpha^2}{24(fK)^{1-\beta}} + \frac{\rho\alpha\nu\beta}{4(fK)^{(1-\beta)/2}} + \frac{2-3\rho^2}{24}\nu^2\right]\varepsilon^2\tau_{ex} + \cdots\right\},$$

through $O(\varepsilon^2)$, where $\zeta$ and $\hat{x}(\zeta)$ are given by B.69b as before. Apart from setting $\varepsilon$ to 1 to recover the original units, this is the formula quoted in section 2, and fitted to the market in section 3.

**B.5. Special cases.** Two special cases are worthy of special treatment: the stochastic normal model ($\beta = 0$) and the stochastic log normal model ($\beta = 1$). Both these models are simple enough that the expansion can be continued through $O(\varepsilon^4)$. For the stochastic *normal* model ($\beta = 0$) the implied volatilities of European calls and puts are

(B.70a) $$\sigma_N(K) = \varepsilon\alpha\left\{1 + \frac{2-3\rho^2}{24}\varepsilon^2\nu^2\tau_{ex} + \cdots\right\}$$

(B.70b) $$\sigma_B(K) = \varepsilon\alpha\frac{\log f/K}{f-K}\cdot\left(\frac{\zeta}{\hat{x}(\zeta)}\right)\cdot\left\{1 + \left[\frac{\alpha^2}{24fK} + \frac{2-3\rho^2}{24}\nu^2\right]\varepsilon^2\tau_{ex} + \cdots\right\}$$

through $O(\varepsilon^4)$, where

(B.70c) $$\zeta = \frac{\nu}{\alpha}\sqrt{fK}\log f/K, \qquad \hat{x}(\zeta) = \log\left(\frac{\sqrt{1-2\rho\zeta+\zeta^2} - \rho + \zeta}{1-\rho}\right).$$

For the stochastic log normal model ($\beta = 1$) the implied volatilities are

(B.71a) $$\sigma_N(K) = \varepsilon\alpha\frac{f-K}{\log f/K}\cdot\left(\frac{\zeta}{\hat{x}(\zeta)}\right)\cdot\left\{1 + \left[-\tfrac{1}{24}\alpha^2 + \tfrac{1}{4}\rho\alpha\nu + \tfrac{1}{24}(2-3\rho^2)\nu^2\right]\varepsilon^2\tau_{ex} + \cdots\right\}$$

$$\sigma_B(K) = \varepsilon\alpha \cdot \left(\frac{\zeta}{\hat{x}(\zeta)}\right) \cdot \left\{1 + \left[\tfrac{1}{4}\rho\alpha\nu + \tfrac{1}{24}(2-3\rho^2)\nu^2\right]\varepsilon^2\tau_{ex} + \cdots\right\} \tag{B.71b}$$

through $O(\varepsilon^4)$, where

$$\zeta = \frac{\nu}{\alpha}\log f/K, \qquad \hat{x}(\zeta) = \log\left(\frac{\sqrt{1-2\rho\zeta+\zeta^2}-\rho+\zeta}{1-\rho}\right). \tag{B.71c}$$

**Appendix C. Analysis of the dynamic SABR model.**

We use *effective medium theory* [24] to extend the preceding analysis to the dynamic SABR model. As before, we take the volatility $\gamma(t)\hat{\alpha}$ and "volvol" $\nu(t)$ to be small, writng $\gamma(t) \longrightarrow \varepsilon\gamma(t)$, and $\nu(t) \longrightarrow \varepsilon\nu(t)$, and analyze

$$d\hat{F} = \varepsilon\gamma(t)\hat{\alpha}C(\hat{F})dW_1, \tag{C.1a}$$

$$d\hat{\alpha} = \varepsilon\nu(t)\hat{\alpha}dW_2, \tag{C.1b}$$

with

$$dW_1dW_2 = \rho(t)dt, \tag{C.1c}$$

in the limit $\varepsilon \ll 1$. We obtain the prices of European options, and from these prices we obtain the implied volatity of these options. After obtaining the results, we replace $\varepsilon\gamma(t) \longrightarrow \gamma(t)$ and $\varepsilon\nu(t) \longrightarrow \nu(t)$ to get the answer in terms of the original variables.

Suppose the economy is in state $\hat{F}(t) = f,\ \hat{\alpha}(t) = \alpha$ at date $t$. Let $V(t, f, \alpha)$ be the value of, say, a European call option with strike $K$ and exercise date $t_{ex}$. As before, define the transition density $p(t, f, \alpha; T, F, A)$ by

$$p(t,f,\alpha;T,F,A)dFdA \equiv \text{prob}\left\{F < \hat{F}(T) < F + dF,\ A < \hat{\alpha}(T) < A + dA \,\middle|\, \hat{F}(t) = f,\ \hat{\alpha}(t) = \alpha\right\} \tag{C.2a}$$

and define

$$P(t,f,\alpha;T,K) = \int_{-\infty}^{\infty} A^2 p(t,f,\alpha;T,K,A)\,dA. \tag{C.2b}$$

Repeating the analysis in Appendix B through equation B.10a, B.10b now shows that the option price is given by

$$V(t,f,a) = [f-K]^+ + \tfrac{1}{2}\varepsilon^2C^2(K)\int_t^{t_{ex}} \gamma^2(T)P(t,f,\alpha;T,K)\,dT, \tag{C.3}$$

where $P(t, f, \alpha; T, K)$ is the solution of the backwards problem

$$P_t + \tfrac{1}{2}\varepsilon^2\left\{\gamma^2\alpha^2C^2(f)P_{ff} + 2\rho\gamma\nu\alpha^2C(f)P_{f\alpha} + \nu^2\alpha^2P_{\alpha\alpha}\right\} = 0, \qquad \text{for } t < T \tag{C.4a}$$

$$P = \alpha^2\delta(f-K), \qquad \text{for } t = T. \tag{C.4b}$$

We eliminate $\gamma(t)$ by defining the new time variable

$$s = \int_0^t \gamma^2(t')dt', \qquad s' = \int_0^T \gamma^2(t')dt', \qquad s_{ex} = \int_0^{t_{ex}} \gamma^2(t')dt'. \tag{C.5}$$

Then the option price becomes

$$V(t,f,a) = [f-K]^+ + \tfrac{1}{2}\varepsilon^2 C^2(K) \int_s^{s_{ex}} P(s,f,\alpha;s',K)\,ds', \tag{C.6}$$

where $P(s,f,\alpha;s',K)$ solves the forward problem

$$P_s + \tfrac{1}{2}\varepsilon^2 \left\{\alpha^2 C^2(f) P_{ff} + 2\eta(s)\alpha^2 C(f) P_{f\alpha} + \upsilon^2(s)\alpha^2 P_{\alpha\alpha}\right\} = 0 \qquad \text{for } s < s' \tag{C.7a}$$

$$P = \alpha^2 \delta(f-K), \qquad \text{for } s = s'. \tag{C.7b}$$

Here

$$\eta(s) = \rho(t)\nu(t)/\gamma(t), \qquad \upsilon(s) = \nu(t)/\gamma(t). \tag{C.8}$$

We solve this problem by using an *effective media* strategy [24]. In this strategy our objective is to determine which constant values $\bar{\eta}$ and $\bar{\upsilon}$ yield the same option price as the the time dependent coefficients $\eta(s)$ and $\upsilon(s)$. If we could find these constant values, this would reduce the problem to the non-dynamic SABR model solved in Appendix B.

We carry out this strategy by applying the same series of *time-independent* transformations that was used to solve the non-dynamic SABR model in Appendix B, defining the transformations in terms of the (as yet unknown) constants $\bar{\eta}$ and $\bar{\upsilon}$. The resulting problem is relatively complex, more complex than the canonical problem obtained in Appedix B. We use a regular perturbation expansion to solve this problem, and once we have solved this problem, we choose $\bar{\eta}$ and $\bar{\upsilon}$ so that all terms arising from the time dependence of $\eta(t)$ and $\upsilon(t)$ cancel out. As we shall see, this simultaneously determines the "effective" parameters and allows us to use the analysis in Appendix B to obtain the implied volatility of the option.

**C.1. Transformation.** As in Appendix B, we change independent variables to

$$z = \frac{1}{\varepsilon\alpha} \int_K^f \frac{df'}{C(f')}, \tag{C.9a}$$

and define

$$B(\varepsilon\alpha z) = C(f). \tag{C.9b}$$

We then change dependent variables from $P$ to $\hat{P}$, and then to $H$:

$$\hat{P} = \frac{\varepsilon}{\alpha} C(K) P, \tag{C.9c}$$

$$H = \sqrt{C(K)/C(f)}\,\hat{P} \equiv \sqrt{B(0)/B(\varepsilon\alpha z)}\,\hat{P}. \tag{C.9d}$$

Following the reasoning in Appendix B, we obtain

$$V(t,f,a) = [f-K]^+ + \tfrac{1}{2}\varepsilon\alpha\sqrt{B(0)B(\varepsilon\alpha z)} \int_s^{s_{ex}} H(s,z,\alpha;s')\,ds', \tag{C.10}$$

where $H(s,z,\alpha;s')$ is the solution of

$$H_s + \tfrac{1}{2}\left(1 - 2\varepsilon\eta z + \varepsilon^2\upsilon^2 z^2\right) H_{zz} - \tfrac{1}{2}\varepsilon^2 \eta\alpha \frac{B'}{B}(zH_z - H) + \varepsilon^2\alpha^2 \left(\tfrac{1}{4}\frac{B''}{B} - \tfrac{3}{8}\frac{B'^2}{B^2}\right) H = 0 \tag{C.11a}$$

for $s < s'$, and

(C.11b) $$H = \delta(z) \qquad \text{at } s = s'$$

through $O(\varepsilon^2)$. See B.29, B.31a, and B.31b. There are no $\alpha$ derivatives in equations C.11a, C.11b, so we can treat $\alpha$ as a parameter instead of a variable. Through $O(\varepsilon^2)$ we can also treat $B'/B$ and $B''/B$ as constants:

(C.12) $$b_1 \equiv \frac{B'(\varepsilon\alpha z_0)}{B(\varepsilon\alpha z_0)}, \qquad b_2 \equiv \frac{B''(\varepsilon\alpha z_0)}{B(\varepsilon\alpha z_0)},$$

where $z_0$ will be chosen later. Thus we must solve

(C.13a) $$H_s + \tfrac{1}{2}\left(1 - 2\varepsilon\eta z + \varepsilon^2 v^2 z^2\right) H_{zz} - \tfrac{1}{2}\varepsilon^2 \eta\alpha b_1 (zH_z - H) + \varepsilon^2\alpha^2 \left(\tfrac{1}{4}b_2 - \tfrac{3}{8}b_1^2\right) H = 0 \qquad \text{for } s < s',$$

(C.13b) $$H = \delta(z) \qquad \text{at } s = s'.$$

At this point we would like to use a time-independent transformation to remove the $zH_z$ term from equation C.13a. It is not possible to cancel this term exactly, since the coefficient $\eta(s)$ is time dependent. Instead we use the transformation

(C.14) $$H = e^{\frac{1}{4}\varepsilon^2\alpha b_1 \delta z^2} \hat{H},$$

where the constant $\delta$ will be chosen later. This transformation yields

$$\hat{H}_s + \tfrac{1}{2}\left(1 - 2\varepsilon\eta z + \varepsilon^2 v^2 z^2\right) \hat{H}_{zz} - \tfrac{1}{2}\varepsilon^2\alpha b_1(\eta - \delta) z\hat{H}_z$$

$$+\tfrac{1}{4}\varepsilon^2\alpha b_1(2\eta + \delta)\hat{H} + \varepsilon^2\alpha^2\left(\tfrac{1}{4}b_2 - \tfrac{3}{8}b_1^2\right)\hat{H} = 0 \qquad \text{for } s < s',$$

(C.15a) $$\hat{H} = \delta(z) \qquad \text{at } s = s',$$

through $O(\varepsilon^2)$. Later the constant $\delta$ will be selected so that the change in the option price caused by the term $\frac{1}{2}\varepsilon^2\alpha b_1\eta z\hat{H}_z$ is exactly offset by the change in price due to $\frac{1}{2}\varepsilon^2\alpha b_1\delta z\hat{H}_z$ term. In this way to the transformation cancels out the $zH_z$ term "on average."

In a similar vein we define

(C.16a) $$I(\varepsilon\bar{v}z) = \sqrt{1 - 2\varepsilon\bar{\eta}z + \varepsilon^2\bar{v}^2 z^2},$$

and

(C.16b) $$x = \frac{1}{\varepsilon\bar{v}}\int_0^{\varepsilon\bar{v}z} \frac{d\zeta}{I(\zeta)} = \frac{1}{\varepsilon\bar{v}}\log\left(\frac{\sqrt{1 - 2\varepsilon\bar{\eta}z + \varepsilon^2\bar{v}^2z^2} - \bar{\eta}/\bar{v} + \varepsilon\bar{v}z}{1 - \bar{\eta}/\bar{v}}\right),$$

where the constants $\bar{\eta}$ and $\bar{v}$ will be chosen later. This yields

(C.17a) $$\hat{H}_s + \tfrac{1}{2}\frac{1 - 2\varepsilon\eta z + \varepsilon^2 v^2 z^2}{1 - 2\varepsilon\bar{\eta}z + \varepsilon^2\bar{v}^2 z^2}(\hat{H}_{xx} - \varepsilon\bar{v}I'(\varepsilon\bar{v}z)\hat{H}_x) - \tfrac{1}{2}\varepsilon^2\alpha b_1(\eta - \delta)x\hat{H}_x$$

$$+\tfrac{1}{4}\varepsilon^2\alpha b_1(2\eta + \delta)\hat{H} + \varepsilon^2\alpha^2\left(\tfrac{1}{4}b_2 - \tfrac{3}{8}b_1^2\right)\hat{H} = 0 \qquad \text{for } s < s',$$

(C.17b) $$\hat{H} = \delta(x) \qquad \text{at } s = s',$$

through $O(\varepsilon^2)$. Here we used $z = x + \cdots$ and $z\hat{H}_z = x\hat{H}_x + \cdots$ to leading order to simplify the results. Finally, we define

(C.18) $$\hat{H} = I^{1/2}(\varepsilon\bar{\upsilon}z)Q.$$

Then the price of our call option is

(C.19) $$V(t,f,a) = [f-K]^+ + \tfrac{1}{2}\varepsilon\alpha\sqrt{B(0)B(\varepsilon\alpha z)}I^{1/2}(\varepsilon\bar{\upsilon}z)e^{\frac{1}{4}\varepsilon^2\alpha b_1\delta z^2}\int_s^{s_{ex}} Q(s,x;s')\,ds',$$

where $Q(s,x;s')$ is the solution of

(C.20a) $$Q_s + \tfrac{1}{2}\frac{1-2\varepsilon\eta z+\varepsilon^2\upsilon^2 z^2}{1-2\varepsilon\bar{\eta}z+\varepsilon^2\bar{\upsilon}^2 z^2}Q_{xx} - \tfrac{1}{2}\varepsilon^2\alpha b_1(\eta-\delta)xQ_x + \tfrac{1}{4}\varepsilon^2\alpha b_1(2\eta+\delta)Q$$
$$+\varepsilon^2\bar{\upsilon}^2\left(\tfrac{1}{4}I''I - \tfrac{1}{8}I'I'\right)Q + \varepsilon^2\alpha^2\left(\tfrac{1}{4}b_2 - \tfrac{3}{8}b_1^2\right)Q = 0 \qquad \text{for } s < s',$$

(C.20b) $$Q = \delta(x) \qquad \text{at } s = s',$$

Using

(C.21) $$z = x - \tfrac{1}{2}\varepsilon\bar{\eta}x^2 + \cdots,$$

we can simplify this to

(C.22a) $$Q_s + \tfrac{1}{2}Q_{xx} = \varepsilon(\eta-\bar{\eta})xQ_{xx} - \tfrac{1}{2}\varepsilon^2\left[\upsilon^2 - \bar{\upsilon}^2 - 3\bar{\eta}(\eta-\bar{\eta})\right]x^2Q_{xx} + \tfrac{1}{2}\varepsilon^2\alpha b_1(\eta-\delta)(xQ_x - Q)$$
$$-\tfrac{3}{4}\varepsilon^2\alpha b_1\delta Q - \varepsilon^2\bar{\upsilon}^2\left(\tfrac{1}{4}I''I - \tfrac{1}{8}I'I'\right)Q - \varepsilon^2\alpha^2\left(\tfrac{1}{4}b_2 - \tfrac{3}{8}b_1^2\right)Q \qquad \text{for } s < s',$$

(C.22b) $$Q = \delta(x) \qquad \text{at } s = s',$$

through $O(\varepsilon^2)$. Note that $I, I'$, and $I''$ can be replaced by the constants $I(\varepsilon\bar{\upsilon}z_0), I'(\varepsilon\bar{\upsilon}z_0)$, and $I''(\varepsilon\bar{\upsilon}z_0)$ through $O(\varepsilon^2)$.

**C.2. Perturbation expansion.** Suppose we were to expand $Q(s,x;s')$ as a power series in $\varepsilon$:

(C.23) $$Q(s,x;s') = Q^{(0)}(s,x;s') + \varepsilon Q^{(1)}(s,x;s') + \varepsilon^2 Q^{(2)}(s,x;s') + \cdots.$$

Substituting this expansion into C.22a, C.22b yields the following hierarchy of equations. To leading order we have

(C.24a) $$Q_s^{(0)} + \tfrac{1}{2}Q_{xx}^{(0)} = 0 \qquad \text{for } s < s',$$

(C.24b) $$Q^{(0)} = \delta(x) \qquad \text{at } s = s'.$$

At $O(\varepsilon)$ we have

(C.25a) $$Q_s^{(1)} + \tfrac{1}{2}Q_{xx}^{(1)} = (\eta-\bar{\eta})xQ_{xx}^{(0)} \qquad \text{for } s < s',$$

(C.25b) $$Q^{(1)} = 0 \qquad \text{at } s = s'.$$

At $O(\varepsilon^2)$ we can break the solution into

(C.26) $$Q^{(2)} = Q^{(2s)} + Q^{(2a)} + Q^{(2b)},$$

where

(C.27a) $$Q_s^{(2s)} + \tfrac{1}{2}Q_{xx}^{(2s)} = -\tfrac{3}{4}ab_1\delta Q^{(0)} - \bar{\upsilon}^2\left(\tfrac{1}{4}I''I - \tfrac{1}{8}I'I'\right)Q^{(0)} - \alpha^2\left(\tfrac{1}{4}b_2 - \tfrac{3}{8}b_1^2\right)Q^{(0)} \qquad \text{for } s < s',$$

(C.27b) $$Q^{(2s)} = 0 \qquad \text{at } s = s',$$

where

(C.28a) $$Q_s^{(2a)} + \tfrac{1}{2}Q_{xx}^{(2a)} = \tfrac{1}{2}\alpha b_1(\eta - \delta)(xQ_x^{(0)} - Q^{(0)}) \qquad \text{for } s < s',$$

(C.28b) $$Q^{(2a)} = 0 \qquad \text{at } s = s',$$

and where

(C.29a) $$Q_s^{(2b)} + \tfrac{1}{2}Q_{xx}^{(2b)} = (\eta - \bar{\eta})xQ_{xx}^{(1)} - \tfrac{1}{2}\left[\upsilon^2 - \bar{\upsilon}^2 - 3\bar{\eta}(\eta - \bar{\eta})\right]x^2Q_{xx}^{(0)} \qquad \text{for } s < s',$$

(C.29b) $$Q^{(2b)} = 0 \qquad \text{at } s = s'.$$

Once we have solved these equations, then the option price is then given by

(C.30a) $$V(t,f,a) = [f-K]^+ + \tfrac{1}{2}\varepsilon\alpha\sqrt{B(0)B(\varepsilon\alpha z)}I^{1/2}(\varepsilon z)e^{\frac{1}{4}\varepsilon^2 ab_1\delta z^2}\,J,$$

where

(C.30b) $$\begin{aligned} J = \int_s^{s_{ex}} Q^{(0)}(s,x;s')\,ds' + \varepsilon\int_s^{s_{ex}} Q^{(1)}(s,x;s')\,ds' + \varepsilon^2\int_s^{s_{ex}} Q^{(2s)}(s,x;s')\,ds' \\ +\varepsilon^2\int_s^{s_{ex}} Q^{(2a)}(s,x;s')\,ds' + \varepsilon^2\int_s^{s_{ex}} Q^{(2b)}(s,x;s')\,ds' + \cdots. \end{aligned}$$

The terms $Q^{(1)}, Q^{(2a)}$, and $Q^{(2b)}$ arise from the time-dependence of the coefficients $\eta(s)$ and $\upsilon(s)$. Indeed, if $\eta(s)$ and $\upsilon(s)$ were constant in time, we would have $Q^{(1)} \equiv Q^{(2a)} \equiv Q^{(2b)} \equiv 0$, and the solution would be just $Q^{(s)} \equiv Q^{(0)} + \varepsilon^2 Q^{(2s)}$. Therefore, we will first solve for $Q^{(1)}, Q^{(2a)}$, and $Q^{(2b)}$, and then *try* to choose the constants $\delta$, $\bar{\eta}$, and $\bar{\upsilon}$ so that the last three integrals are zero for all $x$. In this case, the option price woud be given by

(C.31a) $$V(t,f,a) = [f-K]^+ + \tfrac{1}{2}\varepsilon\alpha\sqrt{B(0)B(\varepsilon\alpha z)}I^{1/2}(\varepsilon z)e^{\frac{1}{4}\varepsilon^2 ab_1\delta z^2}\int_s^{s_{ex}} Q^{(s)}(s,x;s')\,ds',$$

and, through $O(\varepsilon^2)$, $Q^{(s)}$ would be the solution of the static problem

(C.31b) $$Q_s^{(s)} + \tfrac{1}{2}Q_{xx}^{(s)} = -\tfrac{3}{4}\varepsilon^2 ab_1\delta Q^{(s)} - \varepsilon^2\bar{\upsilon}^2\left(\tfrac{1}{4}I''I - \tfrac{1}{8}I'I'\right)Q^{(s)} - \varepsilon^2\alpha^2\left(\tfrac{1}{4}b_2 - \tfrac{3}{8}b_1^2\right)Q^{(s)} \qquad \text{for } s < s',$$

(C.31c) $$Q^{(s)} = \delta(x) \qquad \text{at } s = s'.$$

This is exactly the time-independent problem solved in Appendix B. See equations B.42, B.43a, and B.43b. So *if* we can carry out this strategy, we can obtain option prices for the dynamic SABR model by reducing them to the previously-obtained prices for the static model.

**C.2.1. Leading order analysis.** The solution of C.24a, C.24b is Gaussian:

$$Q^{(0)} = G(x/\sqrt{\Delta}) \tag{C.32a}$$

where

$$G(x/\sqrt{\Delta}) = \frac{1}{\sqrt{2\pi\Delta}} e^{-x^2/2\Delta}, \qquad \Delta = s' - s. \tag{C.32b}$$

For future reference, note that

$$G_x = -\frac{x}{\Delta} G; \qquad G_{xx} = \frac{x^2 - \Delta}{\Delta^2} G; \qquad G_{xxx} = -\frac{x^3 - 3\Delta x}{\Delta^3} G; \tag{C.33a}$$

$$G_{xxxx} = \frac{x^4 - 6\Delta x^2 + 3\Delta^2}{\Delta^4} G; \qquad G_{xxxxx} = -\frac{x^5 - 10\Delta x^3 + 15\Delta^2 x}{\Delta^5} G, \tag{C.33b}$$

$$G_{xxxxxx} = \frac{x^6 - 15\Delta x^4 + 45\Delta^2 x^2 - 15\Delta^3}{\Delta^6} G. \tag{C.33c}$$

**C.2.2. Order $\varepsilon$.** Substituting $Q^{(0)}$ into the equation for $Q^{(1)}$ and using C.33a yields

$$\begin{aligned} Q^{(1)}_s + \tfrac{1}{2} Q^{(1)}_{xx} &= (\eta - \bar{\eta}) \frac{x^3 - \Delta x}{\Delta^2} G \\ &= -(s' - s)(\eta - \bar{\eta}) G_{xxx} - 2(\eta - \bar{\eta}) G_x \qquad \text{for } s < s', \end{aligned} \tag{C.34}$$

with the "initial" condition $Q^{(1)} = 0$ at $s = s'$. The solution is

$$Q^{(1)} = A(s, s') G_{xxx} + 2A_{s'}(s, s') G_x \tag{C.35a}$$

$$= \frac{\partial}{\partial s'} \left\{ 2A(s, s') G_x(x/\sqrt{s' - s} \right\}, \tag{C.35b}$$

where

$$A(s, s') = \int_s^{s'} (s' - \tilde{s})[\eta(\tilde{s}) - \bar{\eta}] d\tilde{s}; \qquad A_{s'}(s, s') = \int_s^{s'} [\eta(\tilde{s}) - \bar{\eta}] d\tilde{s}. \tag{C.35c}$$

This term contributes

$$\int_s^{s_{ex}} Q^{(1)}(s, x; s') ds' = 2A(s, s_{ex}) G_x(x/\sqrt{s_{ex} - s}) \tag{C.36}$$

to the option price. See equations C.30a, C.30b. To eliminate this contribution, we chose $\bar{\eta}$ so that $A(s, s_{ex}) = 0$:

$$\bar{\eta} = \frac{\int_s^{s_{ex}} (s_{ex} - \tilde{s}) \eta(\tilde{s}) d\tilde{s}}{\frac{1}{2}(s_{ex} - s)^2}. \tag{C.37}$$

**C.2.3. The $\varepsilon^2 Q^{(2a)}$ term.** From equation C.28a we obtain

$$\begin{aligned} Q^{(2a)}_s + \tfrac{1}{2} Q^{(2a)}_{xx} &= -\tfrac{1}{2} \alpha b_1 (\eta - \delta) \frac{x^2 + \Delta}{\Delta} G \\ &= -\tfrac{1}{2} \alpha b_1 (\eta - \delta) \Delta G_{xx} - \alpha b_1 (\eta - \delta) G \end{aligned} \tag{C.38}$$

for $s < s'$, with $Q^{(2a)} = 0$ at $s = s'$. Solving then yields

$$Q^{(2a)} = \frac{\partial}{\partial s'} \left\{ \alpha b_1 \int_s^{s'} (s' - \tilde{s})[\eta(\tilde{s}) - \delta] d\tilde{s}\, G(x/\sqrt{s' - s}) \right\}. \tag{C.39}$$

This term makes a contribution of

$$\int_s^{s_{ex}} Q^{(2a)}(s, x; s') ds' = \alpha b_1 \left( \int_s^{s_{ex}} (s_{ex} - \tilde{s})[\eta(\tilde{s}) - \delta] d\tilde{s} \right) G(x/\sqrt{s_{ex} - s}) \tag{C.40}$$

to the option price, so we choose

$$\delta = \bar{\eta} = \frac{\int_s^{s_{ex}} (s_{ex} - \tilde{s})[\eta(\tilde{s}) - \delta] d\tilde{s}}{\frac{1}{2}(s_{ex} - \tilde{s})^2}. \tag{C.41}$$

to eliminate this contribution.

**C.2.4. The $\varepsilon^2 Q^{(2b)}$ term.** Substituting $Q^{(1)}$ and $Q^{(0)}$ into equation C.29a, we obtain

$$Q_s^{(2b)} + \tfrac{1}{2} Q_{xx}^{(2b)} = (\eta - \bar{\eta}) A x G_{xxxxx} + 2(\eta - \bar{\eta}) A_{s'} x G_{xxx} - \tfrac{1}{2} \kappa x^2 G_{xx}, \tag{C.42a}$$

for $s < s'$, where

$$\kappa = \upsilon^2(s) - \bar{\upsilon}^2 - 3\bar{\eta}[\eta(s) - \bar{\eta}]. \tag{C.42b}$$

This can be re-written as

$$\begin{aligned} Q_s^{(2b)} + \tfrac{1}{2} Q_{xx}^{(2b)} = & -(\eta - \bar{\eta}) A [\Delta G_{xxxxxx} + 5 G_{xxxx}] - 2(\eta - \bar{\eta}) A_{s'} [\Delta G_{xxxx} + 3 G_{xx}] \\ & - \tfrac{1}{2} \kappa [\Delta^2 G_{xxxx} + 5 \Delta G_{xx} + 2G] \end{aligned} \tag{C.43}$$

Solving this with the initial condition $Q^{(2b)} = 0$ at $s = s'$ yields

$$\begin{aligned} Q^{(2b)} = & \tfrac{1}{2} A^2(s, s')\, G_{xxxxxx} + 2A(s, s') A_{s'}(s, s')\, G_{xxxx} \\ & + 3 \int_s^{s'} [\eta(\tilde{s}) - \bar{\eta}] A(\tilde{s}, s') d\tilde{s}\, G_{xxxx} + 3A_{s'}^2(s, s')\, G_{xx} \\ & + \tfrac{1}{2} \int_s^{s'} [s' - \tilde{s}]^2 \kappa(\tilde{s}) d\tilde{s}\, G_{xxxx} + \tfrac{5}{2} \int_s^{s'} [s' - \tilde{s}] \kappa(\tilde{s}) d\tilde{s}\, G_{xx} + \int_s^{s'} \kappa(\tilde{s}) d\tilde{s}\, G. \end{aligned} \tag{C.44}$$

This can be written as

$$\begin{aligned} Q^{(2b)} = \frac{\partial}{\partial s'} \Bigg\{ & 4A^2(s, s')\, G_{ss} - 12 \int_s^{s'} [\eta(\tilde{s}) - \bar{\eta}] A(\tilde{s}, s') d\tilde{s}\, G_s \\ & -2 \int_s^{s'} (s' - \tilde{s})^2 \kappa(\tilde{s}) d\tilde{s}\, G_s + \int_s^{s'} (s' - \tilde{s}) \kappa(\tilde{s}) d\tilde{s}\, G \Bigg\} \end{aligned} \tag{C.45}$$

Recall that $\bar{\eta}$ was chosen above so that $A(s, s_{ex}) = 0$. Therefore the contribution of $Q^{(2b)}$ to the option price is

$$\begin{aligned} \int_s^{s_{ex}} Q^{(2b)}(s, x; s') ds' = & -\left( 12 \int_s^{s_{ex}} [\eta(\tilde{s}) - \bar{\eta}] A(\tilde{s}, s_{ex}) d\tilde{s} + 2 \int_s^{s_{ex}} (s_{ex} - \tilde{s})^2 \kappa(\tilde{s}) d\tilde{s} \right) G_s(x/\sqrt{s_{ex} - s}) \\ & + \left( \int_s^{s_{ex}} (s_{ex} - \tilde{s}) \kappa(\tilde{s}) d\tilde{s} \right) G(x/\sqrt{s_{ex} - s}), \end{aligned} \tag{C.46}$$

where $\kappa = \upsilon^2(s) - \bar{\upsilon}^2 - 3\bar{\eta}[\eta(s) - \bar{\eta}]$.

We can choose the remaining "effective media" parameter $\bar{\upsilon}$ to set either the coefficient of $G_s(x/\sqrt{s_{ex}-s})$ or the coefficient of $G(x/\sqrt{s_{ex}-s})$ to zero, but cannot set both to zero to completely eliminate the contribution of the term $Q^{(2b)}$. We choose $\bar{\upsilon}$ to set the coefficient of $G_s(x/\sqrt{s_{ex}-s})$ to zero, for reasons that will become apparent in a moment:

$$\text{(C.47)} \qquad \bar{\upsilon}^2 = \frac{1}{\frac{1}{3}(s_{ex}-\tilde{s})^3}\left\{\int_s^{s_{ex}} (s_{ex}-\tilde{s})^2\upsilon^2(\tilde{s})d\tilde{s} \; - 3\bar{\eta}\int_s^{s_{ex}} (s_{ex}-\tilde{s})^2[\eta(\tilde{s})-\bar{\eta}]d\tilde{s} \right.$$
$$\left. -6\int_s^{s_{ex}}\int_s^{s_1} s_2[\eta(s_1)-\bar{\eta}][\eta(s_2)-\bar{\eta}]ds_2 ds_1. \right.$$

Then the remaining contribution to the option price is

$$\text{(C.48a)} \qquad \int_s^{s_{ex}} Q^{(2b)}(s,x;s')ds' = \tfrac{1}{2}\bar{\kappa}(s_{ex}-s)^2\, G(x/\sqrt{s_{ex}-s}) = \tfrac{1}{2}\bar{\kappa}(s_{ex}-s)^2 Q^{(0)}(s,x;s_{ex}),$$

where

$$\text{(C.48b)} \qquad \bar{\kappa} = \frac{1}{\frac{1}{2}(s_{ex}-s)^2}\int_s^{s_{ex}} (s_{ex}-\tilde{s})[\upsilon^2(\tilde{s}) - \bar{\upsilon}^2]d\tilde{s}.$$

Here we have used $\int_s^{s_{ex}}(s_{ex}-\tilde{s})(\eta(\tilde{s})-\bar{\eta})d\tilde{s} = 0$ to simplify C.48b.

**C.3. Equivalent volatilities.** We can now determine the implied volatility for the dynamic model by mapping the problem back to the static model of Appendix B. Recall from C.30a, C.30b that the value of the option is

$$\text{(C.49a)} \qquad V(t,f,a) = [f-K]^+ + \tfrac{1}{2}\varepsilon\alpha\sqrt{B(0)B(\varepsilon\alpha z)}I^{1/2}(\varepsilon z)e^{\frac{1}{4}\varepsilon^2 ab_1\bar{\eta}z^2}\, J,$$

where

$$\text{(C.49b)} \qquad J = \int_s^{s_{ex}} Q^{(0)}(s,x;s')\,ds' + \varepsilon\int_s^{s_{ex}} Q^{(1)}(s,x;s')\,ds' + \varepsilon^2\int_s^{s_{ex}} Q^{(2s)}(s,x;s')\,ds'$$
$$+\varepsilon^2\int_s^{s_{ex}} Q^{(2a)}(s,x;s')\,ds' + \varepsilon^2\int_s^{s_{ex}} Q^{(2b)}(s,x;s')\,ds' + \cdots,$$

and where we have use $\delta = \bar{\eta}$. We chose the "effective parameters" $\bar{\eta}$ and $\bar{\upsilon}$ so that the integrals of $Q^{(1)}$, $Q^{(2a)}$ contribute nothing to $J$. The integral of $Q^{(2b)}$ then contributed $\frac{1}{2}\varepsilon^2\bar{\kappa}(s_{ex}-s)^2\,Q^{(0)}(s,x;s_{ex})$. The option price is

$$\text{(C.50a)} \qquad J = \int_s^{s_{ex}}\{Q^{(0)}(s,x;s') + \varepsilon^2 Q^{(2s)}(s,x;s')\}\,ds' + .\tfrac{1}{2}\varepsilon^2\bar{\kappa}(s_{ex}-s)^2\,Q^{(0)}(s,x;s_{ex}) + \cdots$$
$$= \int_s^{\hat{s}_{ex}}\{Q^{(0)}(s,x;s') + \varepsilon^2 Q^{(2s)}(s,x;s')\}\,ds' + \cdots$$

through $O(\varepsilon^2)$, where

$$\text{(C.50b)} \qquad \hat{s}_{ex} = s_{ex} + \tfrac{1}{2}\varepsilon^2\bar{\kappa}(s_{ex}-s)^2 + \cdots$$

Through $O(\varepsilon^2)$ we can combine $Q^{(s)} = Q^{(0)}(s,x;s') + \varepsilon^2 Q^{(2s)}(s,x;s')$, where $Q^{(s)}$ solves the static problem

$$\text{(C.51a)} \quad Q_s^{(s)} + \tfrac{1}{2}Q_{xx}^{(s)} = -\tfrac{3}{4}\varepsilon^2 ab_1\delta Q^{(s)} - \varepsilon^2\bar{\upsilon}^2\left(\tfrac{1}{4}I''I - \tfrac{1}{8}I'I'\right)Q^{(s)} - \varepsilon^2\alpha^2\left(\tfrac{1}{4}b_2 - \tfrac{3}{8}b_1^2\right)Q^{(s)} \qquad \text{for } s < s',$$

(C.51b) $$Q^{(s)} = \delta(s-s') \qquad \text{at } s = s',$$

This problem is homogeneous in the time $s$, so its solution $Q^{(s)}$ depends only on the time difference $\tau = s' - s$. The option price is therefore

(C.52) $$V(t,f,a) = [f-K]^+ + \tfrac{1}{2}\varepsilon\alpha\sqrt{B(0)B(\varepsilon\alpha z)}I^{1/2}(\varepsilon z)e^{\frac{1}{4}\varepsilon^2 ab_1\bar{\eta}z^2}\int_0^{\hat{s}_{ex}-s} Q^{(s)}(\tau,x)\,d\tau,$$

where $Q^s(\tau,x)$ is the solution of

(C.53a) $$Q^{(s)}_\tau - \tfrac{1}{2}Q^{(s)}_{xx} = \tfrac{3}{4}\varepsilon^2 ab_1\bar{\eta}Q^{(s)} + \varepsilon^2\bar{\upsilon}^2\left(\tfrac{1}{4}I''I - \tfrac{1}{8}I'I'\right)Q^{(s)} + \varepsilon^2\alpha^2\left(\tfrac{1}{4}b_2 - \tfrac{3}{8}b_1^2\right)Q^{(s)} \qquad \text{for } \tau > 0,$$

(C.53b) $$Q^s = \delta(x) \qquad \text{at } \tau = 0.$$

The option price defined by C.52, C.53a, and C.53b is identical to the static model's option price defined by B.42, B.43a, and B.43b, provided we make the identifications

(C.54) $$\nu \to \bar{\upsilon}, \qquad \rho \to \bar{\eta}/\bar{\upsilon}$$

(C.55) $$\tau_{ex} \to \hat{s}_{ex} - s = s_{ex} - s + \tfrac{1}{2}\varepsilon^2\bar{\kappa}(s_{ex}-s)^2$$

in Appendix B for the original non-dynamic SABR model, provided we make the identifications

(C.56a) $$\tau_{ex} = \tau + \varepsilon^2\int_0^\tau \tilde{\tau}[\upsilon^2(\tilde{\tau}) - \bar{\upsilon}^2]d\tilde{\tau},$$

(C.56b) $$\nu \to \bar{\eta}/\bar{\upsilon}, \qquad \nu \to \bar{\upsilon}.$$

See equations B.42 - B.43b. Following the reasoning in the preceding Appendix now shows that the European call price is given by the formula

(C.57) $$V(t,f,K) = (f-K)\mathcal{N}\left(\frac{f-K}{\sigma_N\sqrt{\tau_{ex}}}\right) + \sigma_N\sqrt{\tau_{ex}}\,\mathcal{G}\left(\frac{f-K}{\sigma_N\sqrt{\tau_{ex}}}\right),$$

with the implied *normal* volatility

(C.58a) $$\sigma_N(K) = \frac{\varepsilon\alpha(f-K)}{\int_K^f \frac{df'}{C(f')}}\cdot\left(\frac{\zeta}{\hat{x}(\zeta)}\right)\cdot \left\{1 + \left[\frac{2\gamma_2-\gamma_1^2}{24}\alpha^2C^2(f_{av}) + \tfrac{1}{4}\bar{\eta}\alpha\gamma_1C(f_{av}) + \frac{2\bar{\upsilon}^2 - 3\bar{\eta}^2}{24} + \tfrac{1}{2}\bar{\theta}\right]\varepsilon^2\tau_{ex} + \cdots\right. .$$

where

(C.58b) $$\zeta = \frac{\bar{\upsilon}}{\alpha}\frac{f-K}{C(f_{av})}, \qquad \hat{x}(\zeta) = \log\left(\frac{\sqrt{1-2\bar{\eta}\zeta/\bar{\upsilon}+\zeta^2} - \bar{\eta}/\bar{\upsilon} + \zeta}{1-\bar{\eta}/\bar{\upsilon}}\right),$$

(C.58c) $$f_{av} = \sqrt{fK}, \qquad \gamma_1 = \frac{C'(f_{av})}{C(f_{av})}, \qquad \gamma_2 = \frac{C''(f_{av})}{C(f_{av})},$$

$$\bar{\theta} = \frac{\int_0^\tau \tilde{\tau}[\upsilon^2(\tilde{\tau}) - \bar{\upsilon}^2]d\tilde{\tau}}{\frac{1}{2}\tau^2}. \tag{C.58d}$$

Equivalently, the option prices are given by Black's formula with the effective Black volatility of

$$\sigma_B(K) = \frac{\alpha \log f/K}{\int_K^f \dfrac{df'}{C(f')}} \cdot \left(\frac{\zeta}{\hat{x}(\zeta)}\right) \cdot \tag{C.59}$$
$$\left\{1 + \left[\frac{2\gamma_2 - \gamma_1^2 + 1/f_{av}^2}{24}\alpha^2 C^2(f_{av}) + \tfrac{1}{4}\bar{\eta}\alpha\gamma_1 C(f_{av}) + \frac{2\bar{\upsilon}^2 - 3\bar{\eta}^2}{24} + \tfrac{1}{2}\bar{\theta}\right]\varepsilon^2\tau_{ex} + \cdots\right.$$

**Appendix D. Analysis of other stochastic vol models.**
Adapt analysis to other SV models. Just quote results?

**Appendix E. Analysis of other stochastic vol models.**
Adapt analysis to other SV models. Just quote results?

## REFERENCES


[1] D.T. Breeden and R. H. Litzenberger, *Prices of state-contingent claims implicit in option prices*, J. Business, 51 (1994), pp. 621-651.
[2] B. Dupire, *Pricing with a smile*, Risk, Jan. 1994, pp. 18–20.
[3] B. Dupire, *Pricing and hedging with smiles*, in Mathematics of Derivative Securities, M.A. H. Dempster and S. R. Pliska, eds., Cambridge University Press, Cambridge, 1997, pp. 103–111
[4] E. DermaN and I. Kani, *Riding on a smile*, Risk, Feb. 1994, pp. 32–39.
[5] E. Derman and I. Kani, *Stochastic implied trees: Arbitrage pricing with stochastic term and strike structure of volatility*, Int J. Theor Appl Finance, 1 (1998), pp. 61–110.
[6] J.M. Harrison and S. Pliska, *Martingales and stochastic integrals in the theory of continuous trading*, Stoch. Proc. Appl, 11 (1981), pp. 215-260.
[7] J.M. Harrison and D. Krebs, *Martingales and arbitrage in multiperiod securities markets*, J. Econ. Theory, 20 (1979), pp. 381-408.
[8] I. Karatzas, J.P. Lehoczky, S.E. Shreve, and G.L. Xus, *Martingale and duality methods for utility maximization in an incomplete market*, SIAM J. Control Optim, 29 (1991), pp. 702–730.
[9] J. Michael Steele, *Stochastic Calculus and Financial Applications*, Springer, 2001
[10] F. Jamshidean, *Libor and swap market models and measures*, Fin. Stoch. 1 (1997), pp. 293-330
[11] F. Black, *The pricing of commodity contracts*, Jour. Pol. Ec., 81 (1976), pp. 167-179.
[12] John C. Hull, *Options, Futures, and Other Derivative Securities*, Prentice Hall, 1997.
[13] P. Wilmott, *Paul Wilmott on Quantitative Finance*, John Wiley & Sons, 2000.
[14] Patrick S. Hagan and Diana E. Woodward, *Equivalent Black volatilities*, App. Math. Finance, **6** (1999), pp. 147–157.
[15] P. S. Hagan, A. Lesniewski and D. E. Woodward, *Geometric optics in finance*, in preparation.
[16] F. Wan, *Mathematical Models and Their Analysis,* Harper-Row, 1989.
[17] J. Hull and A. White, The pricing of options on assets with stochastic volatilities, J. of Finance, **42** (1987), pp. 281-300.
[18] S.L. Heston, *A closed-form solution for options with stochastic volatility with applications to bond and currency options*, Rev of Fin Studies, 6 (1993), pp. 327-343.
[19] A. Lewis, *Option Valuation Under Stochastic Volatility*, Financial Press, 2000.
[20] J.P. Fouque, G. Papanicolaou, K.R. Sirclair, *Derivatives in Financial Markets with Stochastic Volatility*, Cambridge Univ Press, 2000.
[21] N. A. Berner, *Hedging vanna & volga*, DKW, private communicatons
[22] J.D. Cole, *Perturbation Methods in Applied Mathematics*, Ginn-Blaisdell, 1968.
[23] J. Kevorkian and J.D. Cole, Perturbation Methods in Applied Mathematics, Springer-Verlag, 1985.
[24] J.F. Clouets, *Diffusion Approximation of a Transport Process in Random Media*, SIAM J Appl Math, **58** (1998), pp. 1604–1621.
[25] I. Karatzas and S. Shreve, *Brownian Motion and Stochastic Calculus*, Springer, 1988.
[26] B. Okdendal, *Stochastic Differential Equations*, Springer, 1998.
[27] M. Musiela and M. Rutkowski, *Martingale Methods in Financial Modelling*, Springer, 1998.
[28] G. B. Whitham, *Linear and Nonlinear Waves*, Wiley, 1974.
[29] J.C. Neu, *Thesis*, California Institute of Technology, 1978