# Explicit SABR Calibration through Simple Expansions

Fabien Le Floc'h*

* Calypso Technology, 106 rue de La Boétie, 75008 Paris

(v1.0 released June 2014)

Abstract *The SABR stochastic volatility model is a very popular interpolator of implied volatilities, with a given dynamic. This paper presents a simple and very fast method to calibrate the SABR model to given market volatilities, that is to imply the SABR parameters from a given market smile.*

Key Words: stochastic volatility, SABR, calibration, implied volatility, finance

## 1. Introduction

The SABR stochastic volatility model (Hagan *et al.*, 2002) enjoys a high popularity as interpolator of implied volatilities, because, in practice, it fits rather well market implied volatility smiles for interest rate derivatives, fx derivatives and equity derivatives with a few parameters $\alpha, \rho, \nu$, and the dynamic of it can be easily controlled through its $\beta$ parameter, usually defined from historical series analysis for the relevant market.

There are however known shortcomings: it is not arbitrage free as the probability density can become negative for low strikes and long maturities. Many authors have proposed various improvements to the original formula (Obłój, 2008; Johnson and Nonas, 2009; Paulot, 2009; Benaim *et al.*, 2008). More recently the focus has been on finite difference techniques to guarantee the arbitrage-free property (Hagan *et al.*, 2014; Le Floc'h and Kennedy, 2014) in a shifted SABR framework, to allow for negative rates.

In practice, the adjustments to the original formula don't matter much in order to find a good initial guess for the calibration procedure. Once a good initial guess is found, a fast local minimizer like Levenberg-Marquardt can be used to calibrate the specific SABR implementation.

A first calibration method was described in (West, 2005), fitting $\alpha$ exactly to the at-the-money implied volatility, and reducing the problem to a two-dimensional minimization in $(\rho, \nu)$. The Nelder-Mead method is proposed as minimizer. It however still requires some initial guess, and can sometimes be unstable, especially with constraints (Le Floc'h, 2012). A numerical method to find an initial guess that fits exactly the at-the-money volatility and the at-the-money skew is proposed in (Gauthier and Rivaille, 2009). It requires however to solve numerically a two-dimensional non linear system of equations.

In contrast, the method proposed here consists in solving analytically a simple system of equations to fit the at-the-money volatility, skew, and curvature. A similar approach was applied to the Heston model in (Forde *et al.*, 2012) with the difference that, in their case,

*Correspondence Address*: Calypso Technology, 106 rue de La Boétie, 75008 Paris. Email: `fabien_lefloch@calypso.com`

ISSN: Print/ISSN Online/YY/issnum1stp–ppcnt© 2013 Unpublished
DOI:

the Heston parameters describe a full volatility surface and the initial guess procedure relies on two expiries plus the short time at-the-money volatility to solve the 5 Heston parameters exactly. In the SABR case, there is just one expiry to consider and 3 parameters to fit. We will first present our method for the lognormal formula as well as for the normal formula, and then show its accuracy in calibrating real world smiles on the arbitrage-free model from Hagan *et al.* (2014) compared to a global optimization via differential evolution.

## 2. SABR normal and lognormal formulas

### 2.1 The SABR model

In the shifted SABR model, the asset forward follows the following stochastic equation:

$$dF = V(F+b)^{\beta} dW_1 \tag{1}$$

$$dV = \nu V dW_2 \tag{2}$$

with $W_1, W_2$ Brownian motions correlated with correlation $\rho$, and $F(0) = f$, $V(0) = \alpha$. $\nu$ represents the volatility of volatility, $b$ is a displacement allowing for negative rates ($b = 0$ for the classic SABR model).

### 2.2 Normal formula

Given the low rates environment we currently experience, it is now common practice to quote swaptions in terms of normal volatility (bpvol). Taking into account the remarks from Obłój (2008) for the case of the lognormal formula to ensure consistency when $\beta \to 1$, the expansion of the normal volatility adapted from Hagan *et al.* (2002) for the shifted SABR model is:

for $f \neq K$ and $\beta \in [0, 1]$

$$\sigma_N(K) = \frac{f-K}{x(K)} \left[ 1 + \left( g(K) + \frac{1}{4}\rho\nu\alpha\beta(f+b)^{\frac{\beta-1}{2}}(K+b)^{\frac{\beta-1}{2}} + \frac{1}{24}(2-3\rho^2)\nu^2 \right) T \right] \tag{3}$$

with

$$g(K) = \frac{1}{24}(\beta^2 - 2\beta)(f+b)^{\beta-1}(K+b)^{\beta-1}\alpha^2$$

$$\zeta(K) = \frac{\nu}{\alpha(1-\beta)} \left( (f+b)^{1-\beta} - (K+b)^{1-\beta} \right)$$

$$x(K) = \frac{1}{\nu} \log \left( \frac{\sqrt{1 - 2\rho\zeta(K) + \zeta^2(K)} - \rho + \zeta(K)}{1-\rho} \right)$$

When $f = K$,

$$\sigma_N(f) = \alpha(f+b)^{\beta} \left[ 1 + \left( g(f) + \frac{1}{4}\rho\nu\alpha\beta(f+b)^{\beta-1} + \frac{1}{24}(2-3\rho^2)\nu^2 \right) T \right] \tag{4}$$

When $\beta = 1$,

$$g(K) = -\frac{1}{24}\alpha^2 \quad , \quad \zeta(K) = \frac{\nu}{\alpha} \log\left( \frac{f+b}{K+b} \right)$$

When $\beta = 0$,

$$g(K) = 0 \quad , \quad \zeta(K) = \frac{\nu}{\alpha}(f-K)$$

### *2.3 Lognormal formula*

The lognormal formula is very similar to the normal formula, $f - K$ becomes $\log(\frac{f}{K})$ and $g$ includes a small adjustment:

for $f \neq K$ and $\beta \in [0, 1]$,

$$\sigma_B(K) = \frac{1}{x(K)} \log(\frac{f+b}{K+b}) \left[1 + \left(g(K) + \frac{1}{4}\rho\nu\alpha\beta(f+b)^{\frac{\beta-1}{2}}(K+b)^{\frac{\beta-1}{2}} + \frac{1}{24}(2-3\rho^2)\nu^2\right)T\right] \tag{5}$$

with

$$g(K) = \frac{1}{24}(\beta-1)^2(f+b)^{\beta-1}(K+b)^{\beta-1}\alpha^2$$

$$\zeta(K) = \frac{\nu}{\alpha(1-\beta)}\left((f+b)^{1-\beta} - (K+b)^{1-\beta}\right)$$

$$x(K) = \frac{1}{\nu}\log\left(\frac{\sqrt{1-2\rho\zeta(K)+\zeta^2(K)} - \rho + \zeta(K)}{1-\rho}\right)$$

When $f = K$,

$$\sigma_B(f) = \alpha(f+b)^{\beta-1}\left[1 + \left(g(f) + \frac{1}{4}\rho\nu\alpha\beta(f+b)^{\beta-1} + \frac{1}{24}(2-3\rho^2)\nu^2\right)T\right] \tag{6}$$

When $\beta = 1$,

$$g(K) = 0 \quad , \quad \zeta(K) = \frac{\nu}{\alpha}\log\left(\frac{f+b}{K+b}\right)$$

When $\beta = 0$,

$$g(K) = \frac{1}{24(f+b)(K+b)}\alpha^2 \quad , \quad \zeta(K) = \frac{\nu}{\alpha}(f-K)$$

It is important not to forget that Hagan derived the lognormal formula as an approximation of the normal formula (Hagan *et al.*, 2002): it won't be accurate or realistic at all in cases where the implied volatility is very high (see Table 1). This is not so realistic for market implied volatilities, but it can have an influence in the global minimization procedure, as global minimizers like differential evolution will try out extreme values.

Of course, in order to obtain a lognormal volatility, it is also possible to solve with high accuracy for the Black implied volatility of a given normal volatility: we can firstly compute corresponding Bachelier option price with a fast evaluation method (see appendix) and secondly use a robust and accurate Black implied volatility solver (Jäckel, 2013; Li and Lee, 2011).

Table 1.: Black volatility obtained from the direct lognormal approximation compared the one solved from the normal approximation using $\alpha = 3.24, \beta = 1.0, \rho = -0.998, \nu = 1.69, f = 2014, K = f, T = 0.48$.

| Formula | Black volatility | Option price |
|---|---|---|
| Lognormal | 0.9325 | 510.19 |
| Normal | 0.2526 | 140.41 |

For performance reasons, we prefer to rely on the normal formula if the calibration is for bpvols and the lognormal formula if the calibration is for Black volatilities.

### 2.4 *Latest Hagan normal formula*

## 3. Explicit initial guess

### 3.1 *Lognormal volatility*

The idea is to find $\alpha, \rho, \nu$ that matches the at-the-money implied volatility $\sigma_0$, skew $\sigma_0'$ and curvature $\sigma_0''$. A direct application of Hagan formula would lead to a system of trivariate polynomials of degrees 3 and 4, which would then require a numerical method. Instead, we rely a simple lower order expansion of the lognormal formula in the variable $z = \log\left(\frac{K+b}{f+b}\right)$ around $z = 0$ that still captures the smile at-the-money accurately (see Appendix XXX):

$$\sigma_B(z) = \alpha(f+b)^{\beta-1} + \frac{1}{2}\left(\rho\nu - (1-\beta)\alpha(f+b)^{\beta-1}\right) z \\ + \frac{1}{12\alpha(f+b)^{\beta-1}}\left((1-\beta)^2(\alpha(f+b)^{\beta-1})^2 + \nu^2(2-3\rho^2)\right) z^2 \tag{7}$$

This is the same expansion used in (Hagan *et al.*, 2002, equation (3.1a)) to explain the phenomenology and dynamic of the SABR model. We have then:

$$\begin{cases} \sigma_0 &= \alpha(f+b)^{\beta-1} \\ \sigma_0' &= \frac{1}{2}\left(\rho\nu - (1-\beta)\sigma_0\right) \\ \sigma_0'' &= \frac{1}{3\sigma_0}\nu^2 + \frac{1}{6\sigma_0}\left((1-\beta)^2\sigma_0^2 - 3\rho^2\nu^2\right) \end{cases} \tag{8}$$

Note that the derivatives are expressed in log-moneyness towards the variable $z = \log(\frac{K+b}{f+b})$ where $K$ is the strike and $f$ is the forward. This system can be exactly solved to give a first guess $\alpha_0, \rho_0, \nu_0$:

$$\begin{cases} \alpha_0 &= \sigma_0(f+b)^{1-\beta} \\ \nu_0^2 &= 3\sigma_0\sigma_0'' - \frac{1}{2}(1-\beta)^2\sigma_0^2 + \frac{3}{2}\left(2\sigma_0' + (1-\beta)\sigma_0\right)^2 \\ \rho_0 &= \frac{1}{\nu_0}\left(2\sigma_0' + (1-\beta)\sigma_0\right) \end{cases} \tag{9}$$

We make sure to constrain $\nu_0$ to be strictly positive (for example by flooring it to $10^{-4}$) , and $\rho_0$ to be in [-1,1]. We then refine this first guess by solving exactly for the at-the-money volatility with the given $\rho_0$ and $\nu_0$. $\alpha_1$ is the root of the following third order polynomial:

$$\frac{(1-\beta)^2 T}{24(f+b)^{2-2\beta}}\alpha^3 + \frac{\rho\beta\nu T}{4(f+b)^{1-\beta}}\alpha^2 + \left(1 + \frac{2-3\rho^2}{24}\nu^2 T\right)\alpha - \sigma_0(f+b)^{1-\beta} = 0 \tag{10}$$

This is the same polynomial used by West (2005) to reduce the dimension of the calibration problem. $\alpha_1$ is picked as the smallest positive root. In theory, choosing the $\alpha_1$ closest to the previously calibrated $\alpha$ (either from the previous expiry, or from a previous day) could be better. In practice, it turns out to always be the smallest positive root. Our initial guess is then $(\alpha_1, \rho_0, \nu_0)$.

### 3.2 *Normal volatility*

When the input volatilities are normal (bpvol), we can just convert them to equivalent lognormal volatilities (eventually relying on some accurate expansion of the normal volatility at-the-money), and apply the same method as in the previous section to find the initial guess. Instead, we prefer to work directly with normal volatility and derive a simple expansion of the normal formula in $z = \log\left(\frac{K+b}{f+b}\right)$ around $z = 0$ (see Appendix XXX);

$$\sigma_N(z) = \alpha(f+b)^\beta + \frac{1}{2}\left(\rho\nu(f+b) + \beta\alpha(f+b)^\beta\right) z \tag{11}$$

$$+ \left[\frac{1}{12\alpha(f+b)^{\beta-2}}(2\nu^2 - 3\rho^2\nu^2) + \frac{1}{4}\rho\nu(f+b) + \frac{1}{12}(\beta^2+\beta)\alpha(f+b)^\beta\right] z^2 \tag{12}$$

Let $\sigma_0, \sigma_0', \sigma_0''$ be respectively the normal volatility, the slope and the curvature at the money expressed in log-moneyness $z$. We have:

$$\begin{cases} \sigma_0 &= \alpha(f+b)^\beta \\ \sigma_0' &= \frac{1}{2}\rho\nu(f+b) + \frac{1}{2}\beta\sigma_0 \\ \sigma_0'' &= \frac{(f+b)^2}{6\sigma_0}(2\nu^2 - 3\rho^2\nu^2) + \frac{1}{2}\rho\nu(f+b) + \frac{1}{6}(\beta^2+\beta)\sigma_0 \end{cases} \tag{13}$$

This system can be exactly solved to give a first guess $\alpha_0, \rho_0, \nu_0$:

$$\begin{cases} \alpha_0 &= \sigma_0(f+b)^{-\beta} \\ \nu_0^2 &= \frac{1}{(f+b)^2}\left[3\sigma_0\sigma_0'' - \frac{1}{2}(\beta^2+\beta)\sigma_0^2 - 3\sigma_0(\sigma_0' - \frac{1}{2}\beta\sigma_0) + \frac{3}{2}\left(2\sigma_0' - \beta\sigma_0\right)^2\right] \\ \rho_0 &= \frac{1}{\nu_0(f+b)}\left(2\sigma_0' - \beta\sigma_0\right) \end{cases} \tag{14}$$

As for the lognormal volatility guess, we refine this first guess by solving exactly for the at-the-money volatility with the given $\rho_0$ and $\nu_0$. $\alpha_1$ is a root of the following third order polynomial:

$$\frac{(\beta^2-2\beta)T}{24(f+b)^{2-2\beta}}\alpha^3 + \frac{\rho\beta\nu T}{4(f+b)^{1-\beta}}\alpha^2 + \left(1 + \frac{2-3\rho^2}{24}\nu^2 T\right)\alpha - \sigma_0(f+b)^{-\beta} = 0 \tag{15}$$

Note that this polynomial is slightly different from the lognormal volatility polynomial. We take the $\alpha_1$ to be the smallest positive root.

### 3.3 *How to find the at-the-money volatility, slope and curvature?*

The simplest is to fit a parabola to the three closest points around the forward with coordinates $(z_{-1}, \hat\sigma_{-1}), (z_0, \hat\sigma_0), (z_1, \hat\sigma_1)$. This is equivalent to a 3 points finite difference on a non uniform grid. We have then:

$$\sigma_0 = z_0 z_1 w_{-1}\hat\sigma_{-1} + z_{-1}z_1 w_0\hat\sigma_0 + z_{-1}z_0 w_1\hat\sigma_1 \tag{16}$$

$$\sigma_0' = -(z_0+z_1)w_{-1}\hat\sigma_{-1} - (z_{-1}+z_1)w_0\hat\sigma_0 - (z_{-1}+z_0)w_1\hat\sigma_1 \tag{17}$$

$$\sigma_0'' = 2w_{-1}\hat\sigma_{-1} + 2w_0\hat\sigma_0 + 2w_1\hat\sigma_1 \tag{18}$$

with

$$w_{-1} = \frac{1}{(z_{-1}-z_0)(z_{-1}-z_1)} \tag{19}$$

$$w_0 = \frac{1}{(z_0-z_{-1})(z_0-z_1)} \tag{20}$$

$$w_1 = \frac{1}{(z_1-z_{-1})(z_1-z_0)} \tag{21}$$

In our experience, using higher order approximations like a 5 points finite difference (equivalent to a quartic on 5 points), or a cubic spline does not lead to any visible improvement in accuracy on our problem.

When the input data is noisy, one could resort to a best fit parabola around the five closest points from the forward, which can be reduced to solving simple linear system. When the volatility at the forward is already present in the input data, we make sure the parabola passes exactly through it.

Another approach is to repeat the initial guess procedure with a parabola on different sets of three points further away from the forward and select the guess corresponding to the best overall fit.

In our numerical tests we will select the guess that gives the minimum mean square error in volatilities between the two guesses stemming from the 3-points parabola and the 5-points parabola.

## 4. Almost exact inversion of SABR smiles

We use the SABR parameters found by calibration on the S&P500 options in December 2008 as starting point (Table 2). From those we regenerate a discrete set of implied volatilities for 12 strikes and 11 expiries using the lognormal SABR formula with $\beta = 1$. And finally we apply our initial guess procedure to each expiry.

Table 2.: Reference SABR parameters

| Expiry | $\alpha$ | $\rho$ | $\nu$ |
|---|---|---|---|
| 0.058 | 0.271 | -0.345 | 1.010 |
| 0.153 | 0.256 | -0.321 | 0.933 |
| 0.230 | 0.256 | -0.346 | 0.820 |
| 0.479 | 0.255 | -0.370 | 0.629 |
| 0.729 | 0.257 | -0.403 | 0.528 |
| 1.227 | 0.260 | -0.429 | 0.448 |
| 1.726 | 0.261 | -0.440 | 0.392 |
| 2.244 | 0.262 | -0.445 | 0.355 |
| 2.742 | 0.262 | -0.445 | 0.329 |
| 3.241 | 0.262 | -0.447 | 0.310 |
| 4.239 | 0.263 | -0.452 | 0.284 |

The root mean square error in implied volatilities of the fit is always lower than $3 \cdot 10^{-4}$. $\rho, \nu$ are recovered with an accuracy higher than $5 \cdot 10^{-3}$ and $\alpha$ is recovered with and accuracy higher than $10^{-4}$, independently of the expiry (see Figure 1).

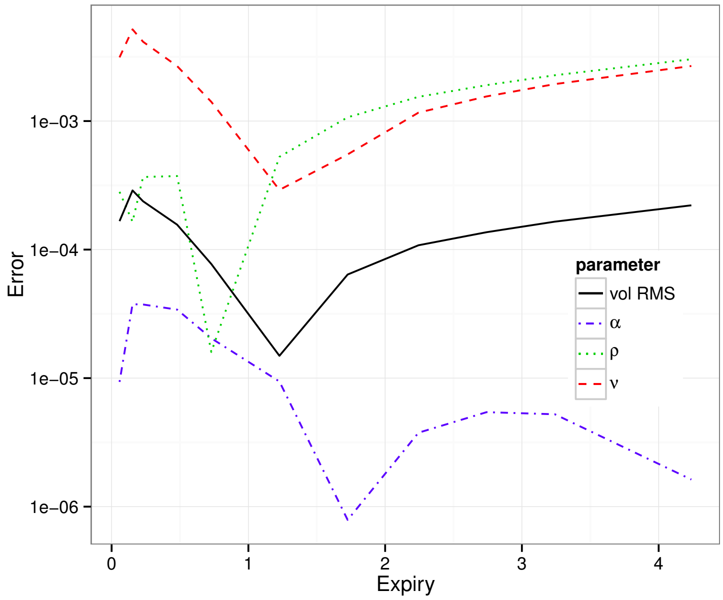
## 5. Real world smiles

### 5.1 *Calibration on the analytic formula*

#### 5.1.1 *Equity option smiles*

To calibrate SABR to equity implied volatility surfaces, we minimize the root mean square error in implied volatilities with the Levenberg-Marquardt method (), using as initial guess, either our explicit method, or alternatively a guess found by running the differential evolution algorithm () with a population size 20 on 1000 generations. The resulting mean square error is displayed in Figure 2.

Figure 1.: Error between the initial guess and the original SABR parameters




The explicit guess leads a mean square error in implied volatilities nearly indistinguishable from the differential evolution guess. It is interesting however to look at the calibrated parameters values: the calibrated $\alpha$ shows some strong variations with the guess found by differential evolution, while it is relatively smooth with our explicit initial guess (Figure 3). $\rho$ and $\nu$ behave similarly. We will take a closer look at why this happens in the next section.

#### *5.1.2 Two SABRs for the same smile*

The fact that two different SABR parameters sets can lead to smiles with nearly the same error measure is the root of issues with differential evolution: it is not obvious which set of parameter is the right one without some additional constraint on the variation of the parameters. This can be resolved by some penalty term, but it is difficult in practice to find the correct penalty term that works for a variety of situations. A better approach is to look beyond the best candidate, and consider as alternative initial guesses the other candidates in the latest generation with very similar error measure and different (lower) $\alpha$.

#### *5.1.3 Swaption smiles*

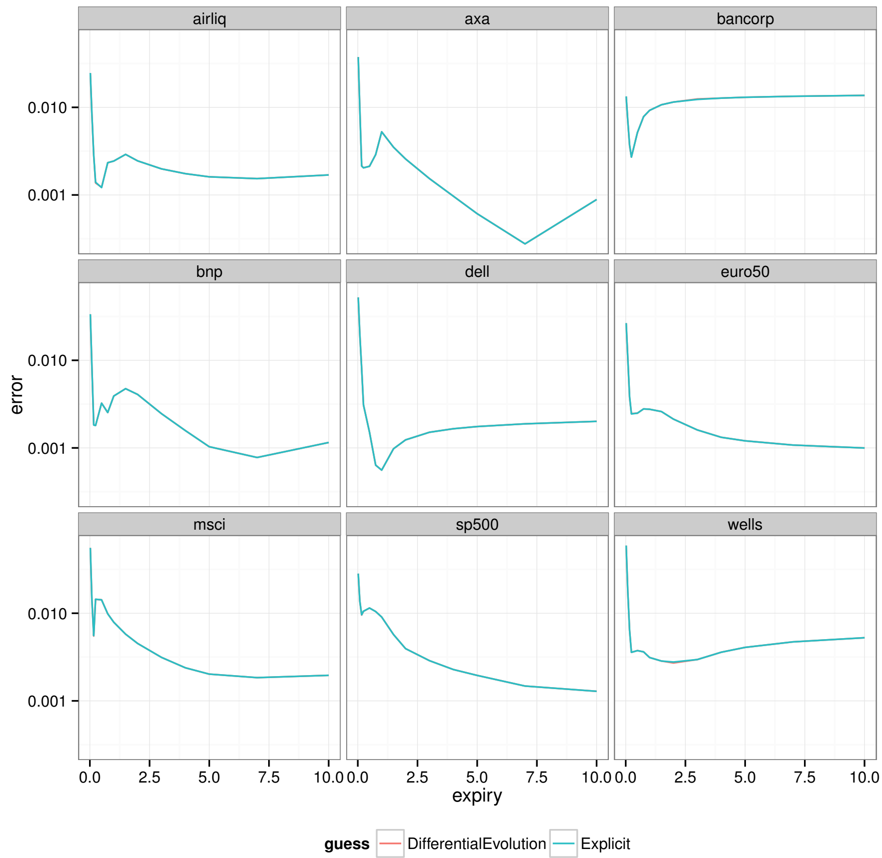
We apply the same calibration methodology as for the equity smiles cases, but relying on the normal formula with $\beta = \frac{1}{2}$ instead. The calibrated parameters are indistinguishable for between the calibration on the explicit initial guess and the calibration on the guess found by differential evolution (see Table XXX). Furthermore, the initial guess is very close to the calibration result.

Figures 4 and 5 show how close are the smile produced by the initial guess and the smile resulting from the Levenberg-Marquardt minimization.

### *5.2 Calibration on the arbitrage free PDE*

## 6. Conclusion

Figure 2.: Root mean square error of calibration on various equity surfaces




## References

Benaim, S., Dodgson, M. and Kainth, D., An arbitrage-free method for smile extrapolation. (2008) , Technical report, Working Paper, QuaRC, Royal Bank of Scotland.

Forde, M., Jacquier, A. and Lee, R. (2012) The small-time smile and term structure of implied volatility under the Heston model, *SIAM Journal on Financial Mathematics*, 3(1), pp. 690–708.

Gauthier, P. and Rivaille, P. Y. H. (2009) Fitting the smile, Smart parameters for SABR and Heston, *Smart Parameters for SABR and Heston (October 30, 2009)*.

Hagan, P. S., Kumar, D., Lesniewski, A. S. and Woodward, D. E. (2002) Managing smile risk, *Wilmott magazine*.

Hagan, P. S., Kumar, D., Lesniewski, A. S. and Woodward, D. E. (2014) Arbitrage free SABR, *Wilmott magazine*.

Jäckel, P. (2013) Lets be rational. , `http://www.pjaeckel.webspace.virginmedia.com/LetsBeRational.pdf` (accessed 03-April-2014).

Johnson, S. and Nonas, B. (2009) Arbitrage-free construction of the swaption cube, *Wilmott Journal*, 1(3), pp. 137–143.

Le Floc'h, F. (2012) Issues of Nelder-Mead Simplex Optimisation with Constraints, *SSRN eLibrary* `http://papers.ssrn.com/abstract=2097904`.

Le Floc'h, F. and Kennedy, G. (2014) Finite difference techniques for arbitrage free SABR, *SSRN eLibrary* `http://papers.ssrn.com/abstract=2402001`.

Li, M. and Lee, K. (2011) An adaptive successive over-relaxation method for computing the Black–Scholes implied volatility, *Quantitative Finance*, 11(8), pp. 1245–1269.

Lorig, M., Pagliarani, S. and Pascucci, A. (2014) Explicit implied vols for multifactor local-stochastic vol models, *arXiv preprint arXiv:1306.5447v3*.

Oblój, J. (2008) Fine-tune your smile: Correction to Hagan et al, *Wilmott Magazine*.

Paulot, L. (2009) Asymptotic implied volatility at the second order with application to the SABR model, *Available at SSRN 1413649*.

West, G. (2005) Calibration of the SABR model in illiquid markets, *Applied Mathematical Finance*, 12(4), pp. 371–385.

Figure 3.: Calibrated $\alpha$ on various equity surfaces

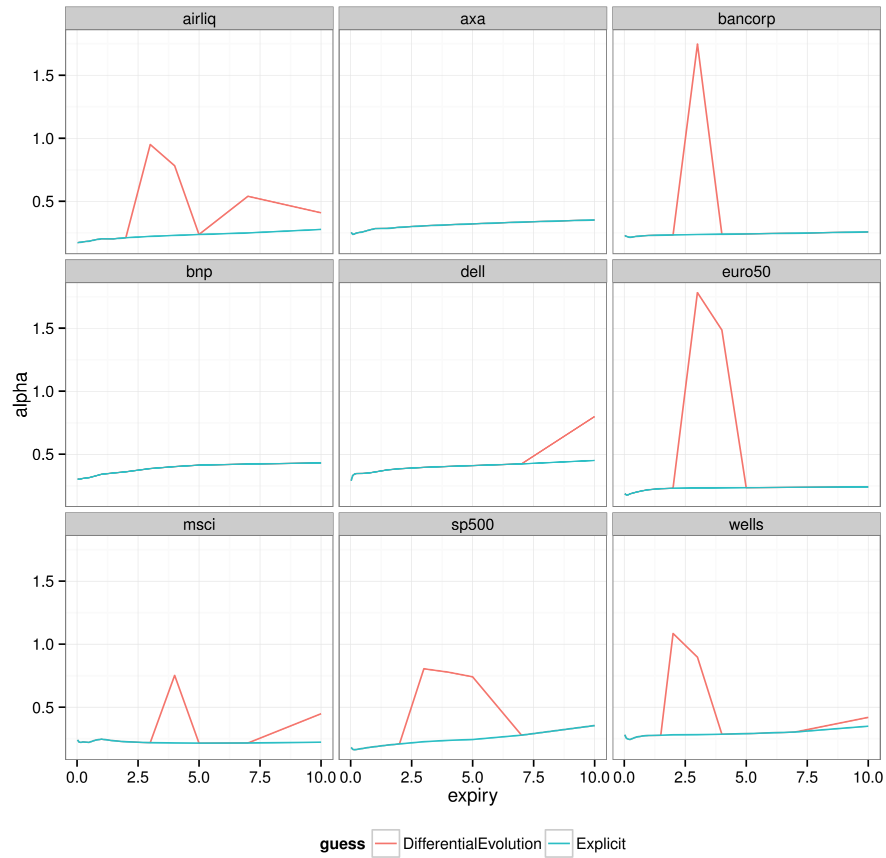



Figure 4.: Initial guess and calibrated smile for a May 2014 1m5y Swaption

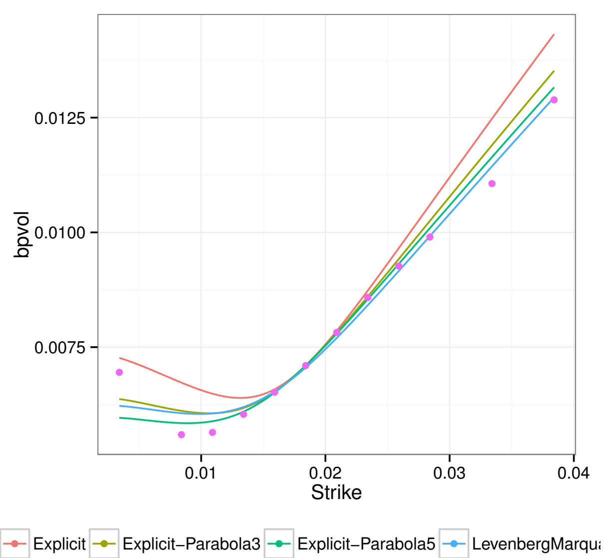


Figure 5.: Initial guess and calibrated smile for a May 2014 2y5y Swaption