

# Calibration Methods of Hull-White Model

Sébastien Gurrieri<sup>1\*</sup>, Masaki Nakabayashi<sup>1§</sup> and Tony Wong<sup>1¶</sup>

<sup>1</sup>*Risk Management Department, Mizuho Securities*  
Tokyo

## Abstract

We describe several strategies for the calibration of one factor Hull-White model with constant or time-dependent mean reversion and volatility parameters to the interest rate vanillas. We propose an efficient approximation formula for the swaption implied volatility which enables us to estimate the mean reversion independently of the volatility. We give the closed-forms for exact pricing using explicit integrals of the model parameters and propose parametric forms for the mean reversion and volatility. We test their performance in terms of quality of fitting and stability w.r.t. market changes, and show that excellent fits can be obtained without suffering from instabilities. Furthermore, our calibration methods and parameter control techniques allow for an elegant interpretation of market moves, which we illustrate with an in-depth analysis of Lehman crisis in the fall of 2008.

---

\*email: [sebastien.gurrieri@mizuho-sc.com](mailto:sebastien.gurrieri@mizuho-sc.com)

§email: [masaki.nakabayashi@mizuho-sc.com](mailto:masaki.nakabayashi@mizuho-sc.com)

¶email: [shekkeung.wong@mizuho-sc.com](mailto:shekkeung.wong@mizuho-sc.com)

# Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Model Dynamics and Closed Forms</b>	<b>4</b>
2.1	Dynamics . . . . .	4
2.2	Closed Forms for Caplets and Swaptions . . . . .	5
2.3	An Approximation of the Swaption Implied Volatility . . . . .	5
2.4	Parameter Specifications . . . . .	7
<b>3</b>	<b>Calibration Strategies</b>	<b>9</b>
3.1	Relations Market/Model parameters . . . . .	9
3.2	Optimizing on the Mean Reversion or not . . . . .	11
3.3	Bootstrap vs. Overall Calibration . . . . .	12
3.4	Calibration to Swaptions . . . . .	12
3.4.1	Method 1 . . . . .	13
3.4.2	Method 2 . . . . .	13
3.4.3	Method 3 . . . . .	14
<b>4</b>	<b>Numerical Results</b>	<b>15</b>
4.1	Performance of SMM approximation . . . . .	15
4.2	Choice of Instruments . . . . .	16
4.3	Co-terminal 20Y . . . . .	16
4.3.1	Conclusion . . . . .	19
4.4	Co-terminals 10Y and 20Y . . . . .	19
4.4.1	Constant $a$ and $\sigma$ . . . . .	19
4.4.2	Constant $a$ and Time-Dependent $\sigma(t)$ . . . . .	21
4.4.3	Time-Dependent $a(t)$ and $\sigma(t)$ . . . . .	24
4.4.4	Conclusion . . . . .	24
4.5	All Instruments . . . . .	24
4.5.1	Conclusion . . . . .	29
4.6	Model and Market Evolution: analysis at Lehman Crisis . . . . .	29
4.7	Bermudan Swaptions . . . . .	31
<b>5</b>	<b>Conclusion</b>	<b>33</b>
<b>A</b>	<b>Integrals</b>	<b>34</b>
<b>B</b>	<b>Simulated Annealing</b>	<b>36</b>
<b>C</b>	<b>Other Numerical Results</b>	<b>37</b>
C.1	SMM vs. Jamshidian Method . . . . .	37
C.2	Other JPY calibration results . . . . .	38
C.3	Other currencies: USD . . . . .	40
C.4	Other currencies: EUR . . . . .	41

# 1 Introduction

Market models, pioneered by [2], [9], have recently emerged as a market standard for the pricing of exotic interest rate products. In these models, the dynamics of observable market rates, such as Libor and swap rates, are directly specified. They are appealing to market practitioners largely due to the fact that they can be calibrated to market caplet and swaption volatilities as well as an initial yield curve. In spite of their increasing popularity, market models have one serious drawback. An accurate implementation can only be made through simulation, which is typically slow due to the large number of state variables (i.e. the discrete-tenored market rates) needed to be evolved through time, and the complicated drift terms associated with the underlying stochastic differential equations (SDEs). This can be a problem for path dependent payoffs which require a large number of simulated paths in order to obtain sufficiently accurate price and risk figures. The problem can be even more acute for products with an early exercise provision as simulation is not naturally suited for performing backward-in-time calculations needed to determine the optimal exercise strategy.

In contrast to market models, affine term structure models (ATSMs) attempt to model the unobservable short rate, assumed to be an affine function of some latent factors. ATSMs have a few appealing properties. First, they are analytically tractable. In particular, closed-form solutions for caps/floors and efficient price approximation methods for European swaptions are often available. Second, Monte Carlo implementation of ATSMs is relatively straightforward compared to market models. In low dimensional cases, path-independent products with an early exercise provision can also be evaluated efficiently using lattice methods. Third, under ATSMs, all kinds of interest rates (e.g. forward Libor and swap rates with any expiry/maturity/payment dates) can be computed readily from the short rate factors, whereas in market models interpolation and extrapolation are often needed when the product dates do not align with the canonical model dates. The computational efficiency offered by ATSMs explains why many banks still use these models for risk calculation and other risk management purposes even after the introduction of the market models.

The main objective of this paper is to analyze in details the issues of model calibration and parameter control associated with a popular ATSM, known as one-factor Hull White (HW1F) model ([5], see [3] for a review). The HW1F model, which is characterized by a mean reversion and a volatility parameter, has been and is still very popular among market practitioners because of its parsimony and analytical tractability (thanks to the normality of the short rate). In spite of its popularity, the existing studies on the model calibration, especially in the case of time-dependent parameters, are rather scarce. In [6] and [7], the authors have covered the basic procedures for calibrating the HW1F model utilizing tree methods. However, there does remain several open issues related to its practical use.

First, it is frequently claimed that adopting time-dependent parameters for the HW1F model can introduce over-parameterization. Yet, numerical examples backing such a statement are rare in the existing literature. As a result, it is unclear in what contexts and to what extent such a statement can be justified. This also raises the question of how one should choose the model parameter configuration, given a particular application at hand. Another challenging issue is related to the calibration, especially in cases where time-dependent parameters are used. Given a certain parameter configuration, a natural question arising is what strategy one should use in calibrating the model. For example, one may ask what parameter constraints should be put in place during the calibration or if the model has some separable property that can make the calibration more efficient. Finally, from traders' perspectives, one of the most important issues lies in control of the model parameters. The reason is that traders often want to incorporate their particular views on the future market moves into the pricing model through the manipulation of the model parameters. This translates into the requirement of understanding the pricing impacts of the model parameters.

In this work, we address the issues mentioned in the previous paragraph. This study provides an explicit realization of some ideas developed in [6] and [7] and supplements them with a thorough analysis of various issues involved such as parameter stability and relationship between model parameters and market prices of calibrating instruments. Hence, our work can be considered as a complement and a detailed follow-up for [6] and [7].

We introduce parametric forms for the time-dependent parameters, with additional constraints. These give us control over the parameters while still allowing enough fitting freedom. We describe several calibration strategies, namely local and global, with constant and/or time-dependent parameters, with separate or simultaneous optimizations. We propose an approximation of the swaption implied volatilities, which, in the particular case of constant mean reversion and volatility, leads to a very efficient (quasi-instantaneous) calibration method. This approximation, due to its simplicity and explicit form, provides further insights into the market/model relationships. The case of Lehman shock in the fall of 2008 and its consequences on the parameters are analysed in details. We emphasize the importance of the mean reversion parameter when attempting to fit over several different tenors, possibly over the whole swaption matrix. We finally conclude that, contrary to common beliefs, excellent fits to the whole swaption matrix can be obtained without traces of instabilities, when both the mean reversion and volatility are time-dependent.

The remainder of this paper is organized as follows. In Section 2, we provide a brief review of the HW1F model. Some analytical formulas relevant to our analysis are provided. In particular, we present an analytical approximation for implied swaption volatilities. In Section 3, we propose a few strategies for calibrating the model under various parameter configurations. In Section 4, we present numerical results to test all the calibration strategies presented previously, and discuss their performance. Finally, we conclude in Section 5.

## 2 Model Dynamics and Closed Forms

Hull-White model definition and properties are well-known and therefore we recall them only briefly here. Our main purpose is instead to provide the notations used in the closed-form calculations with the view to calibrating the model to vanilla options. More precisely, we want to display the variance of a zero-coupon bond ratio and the expressions of the functions characterizing the affine structure of the bond, in order to price swaptions analytically.

The closed-forms can be written thanks to various integrals of the mean reversion and volatility that can be found in [4]. We recall them in appendix A and write them slightly more explicitly, together with additional properties. We keep them under a symbolic form as long as possible such that the formulae remain valid for any specific parametric form that the user may choose. Note that they become explicit as soon as the mean reversion and volatility parametric expressions are specified. In particular, when these are constant, the integrals reduce to their standard form, given for instance in [3].

### 2.1 Dynamics

In the risk neutral measure, denoted by  $\mathbb{Q}$ , the Hull-White short-rate SDE reads

$$dr(t) = \left(\theta(t) - a(t)r(t)\right)dt + \sigma(t)dW^{\mathbb{Q}}(t), \quad (1)$$

with the time-dependent functions  $a(t)$  for the mean reversion and  $\sigma(t)$  for the volatility. Exact replication of the initial curve is realized by fixing the function  $\theta(t)$  to the expression (39) given in appendix A.  $r(t)$  has a normal distribution with mean and variance

$$E[r(t) | \mathcal{F}_s] = \frac{E(s)}{E(t)}r(s) + \alpha(t) - \frac{E(s)}{E(t)}\alpha(s) \quad (2)$$

$$\text{Var}[r(t) | \mathcal{F}_s] = V_r(s, t) \quad (3)$$

where  $\mathcal{F}_s$  is a  $\sigma$ -field capturing the information generated by the process  $r(t)$  up to time  $s$ ,  $E(t)$  is given in (30),  $\alpha(t)$  in (36) and  $V_r(s, t)$  in (37).

The zero-coupon bond  $P(t, T)$  has the affine structure

$$P(t, T) = \exp\left(A(t, T) - B(t, T)r(t)\right) \quad (4)$$

where the functions  $A(t, T)$  and  $B(t, T)$  are given in (43) and (31). It is log-normal with SDE

$$dP(t, T)/P(t, T) = r(t)dt - \sigma(t)B(t, T)dW^{\mathbb{Q}}(t) \quad (5)$$

To calculate closed forms, we are particularly interested in the bond ratio with fixing and paying times  $T_F$  and  $T_P$  ( $t \leq T_F \leq T_P$ ), which has the dynamics in the  $T_P$ -forward measure

$$d\frac{P(t, T_F)}{P(t, T_P)} = \frac{P(t, T_F)}{P(t, T_P)}\sigma(t)\left(B(t, T_P) - B(t, T_F)\right)dW^{T_P}(t) \quad (6)$$

with integrated variance

$$V_p(t, T_F, T_P) = \int_t^{T_F} \sigma^2(u) \left( B(u, T_P) - B(u, T_F) \right)^2 du \quad (7)$$

$$= V_r(t, T_F) B(T_F, T_P)^2. \quad (8)$$

Equation (8) holds due to property (34).

## 2.2 Closed Forms for Caplets and Swaptions

### Caplets

Let us consider a caplet with strike  $K$ , fixing time  $T_F$  and paying time  $T_P$ . This can be rewritten as a zero-bond put option (ZBP), priced by Black Formula with the variance of the bond ratio in (8), i.e.

$$\text{Caplet}(K, T_F, T_P) = (1 + K\delta) \text{ZBP}(T_F, T_P, \frac{1}{1 + K\delta}) \quad (9)$$

$$\text{ZBP}(T_F, T_P, X) = X P(0, T_F) \mathcal{N}(d_+) - P(0, T_P) \mathcal{N}(d_-) \quad (10)$$

$$d_{\pm} = \frac{\ln\left(\frac{P(0, T_F) X}{P(0, T_P)}\right)}{\sqrt{V_p(0, T_F, T_P)}} \pm \frac{1}{2} \sqrt{V_p(0, T_F, T_P)} \quad (11)$$

where  $\delta = \delta(T_F, T_P)$  is the accrual between  $T_F$  and  $T_P$ .

### Swaptions

Let us consider a (payer) swaption with strike  $K$ , maturity  $T_0$ , swap tenor  $T_P$ , and swap cash-flow times  $\{T_i\}_{i=1..n}$ , with  $T_n = T_P$ . This can be rewritten as a weighed sum of zero-bond (put) options using Jamshidian's decomposition [8]. More precisely,

$$\text{PSwaption}(K, T_0, T_P) = \sum_{i=1}^n c_i \text{ZBP}(T_0, T_i, X_i) \quad (12)$$

$$c_i = K\delta(T_{i-1}, T_i) \quad i = 1..n-1 \quad (13)$$

$$c_n = 1 + K\delta(T_{n-1}, T_n) \quad (14)$$

$$X_i = \exp\left(A(T_0, T_i) - B(T_0, T_i)r^*\right) \quad (15)$$

where  $r^*$  satisfies the equation

$$\sum_{i=1}^n c_i \exp\left(A(T_0, T_i) - B(T_0, T_i)r^*\right) = 1. \quad (16)$$

## 2.3 An Approximation of the Swaption Implied Volatility

Although (12) gives us an analytical expression of the swaption price, it is not explicit in the model parameters, and the dependence of the implied volatility derived from the price  $\text{PSwaption}(K, T_0, T_P)$  on  $a(t)$  and  $\sigma(t)$  is far from obvious. For calibration purposes, it may be desirable to see a more

direct relation between the market implied volatility and the model parameters. In this section, we suggest an approximation satisfying this goal.

From the definition of swaptions it is clear that if we could calculate the variance of a swap rate having a log-normal distribution (i.e. the SMM model), we would be able to calibrate simply by equating this variance with that obtained from the market implied volatility. Unfortunately, Hull-White model is not a swap rate model, the swap rate is not log-normal. However the bond ratio  $\frac{P(t, T_F)}{P(t, T_P)}$  is. This observation directs us toward an approximation relating the swap rate to the bond ratio above. We follow the strategy proposed in [3] for BGM model and adapt it to the Hull-White case.

Denote by  $S(t, T_0, T_n)$  the swap rate prevailing at time  $t$  for the swap starting at  $T_0$  and ending on  $T_n$ , with  $t < T_0 < T_n$ . It is defined by

$$S(t, T_0, T_n) = \frac{P(t, T_0) - P(t, T_n)}{\sum_{i=1}^n \delta(T_{i-1}, T_i) P(t, T_i)}. \quad (17)$$

The approximation consists in assuming it is log-normal under the annuity measure  $\mathcal{A}$  and try to calculate its variance, the quantity we need for pricing. To this end, we approximate the true swap rate  $S(t, T_0, T_n)$  by  $\tilde{S}(t, T_0, T_n)$  with

$$\tilde{S}(t, T_0, T_n) = \frac{P(0, T_n)}{\sum_{i=1}^n \delta(T_{i-1}, T_i) P(0, T_i)} \left[ \frac{P(t, T_0)}{P(t, T_n)} - 1 \right]. \quad (18)$$

Now applying Ito's Lemma to the approximated swap rate above, we find

$$\begin{aligned} \frac{d\tilde{S}(t, T_0, T_n)}{\tilde{S}(t, T_0, T_n)} &= \frac{1}{\tilde{S}(t, T_0, T_n)} \frac{P(0, T_n)}{\sum_{i=1}^n \delta(T_{i-1}, T_i) P(0, T_i)} d \left( \frac{P(t, T_0)}{P(t, T_n)} \right) \\ &= \frac{S(0, T_0, T_n) P(0, T_n) P(t, T_0)}{\tilde{S}(t, T_0, T_n) [P(0, T_0) - P(0, T_n)] P(t, T_n)} \sigma(t) [B(t, T_n) - B(t, T_0)] dW^{T_n}(t) \\ &\approx \text{drift} + \frac{P(0, T_0)}{P(0, T_0) - P(0, T_n)} \sigma(t) (B(t, T_n) - B(t, T_0)) dW^{\mathcal{A}}(t). \end{aligned}$$

In reaching the final line above, we have replaced  $\tilde{S}(t, T_0, T_n)$  and  $\frac{P(t, T_0)}{P(t, T_n)}$  by their initial values and changed measure from the  $T_n$ -forward to the annuity measure  $\mathcal{A}$ . This introduces a drift but for pricing purpose we are only interested in the variance, which reads

$$V_{\text{swap}}(T_0, T_n) = \left[ \frac{P(0, T_0)}{P(0, T_0) - P(0, T_n)} \right]^2 V_p(0, T_0, T_n) \quad (19)$$

where  $V_p(0, T_0, T_n)$  is the variance of the bond ratio as defined in formula (7).

Beware that this approximation is crude and should not be considered as a replacement of (12) for analytical pricing. Its quality is investigated numerically in chapter 4. Let us point out that we

intend to use it only to estimate the mean reversion, and later correct this estimation by a fitting to the actual analytical price.

Finally, note that due to the specific form of  $V_p(0, T_0, T_n)$  in (8), the Hull-White implied variance  $V_{\text{swap}}(T_0, T_n)$  depends on the model volatility  $\sigma(t)$  only through the function  $V_r(0, T_0)$ , in other words the dependences in the swap tenor  $T_n$  and  $\sigma(t)$  are decoupled. This means that the ratio of two implied volatilities having the same maturity  $T_0$  is independent of  $\sigma(t)$ , a feature that will enable us to estimate the mean reversion without knowledge of the volatility.

## 2.4 Parameter Specifications

One way to achieve at the same time ease of implementation and ability to integrate the functions  $a(t)$  and  $\sigma(t)$  analytically is to set these to piecewise constant functions. This is what we will always do from now on, constant  $a$  and  $\sigma$  being just particular cases.

Let us call the parameter time grid  $\{t_i\}_{i=0..N-1}$  where  $N$  is the number of points on the grid. Since  $a(t)$  and  $\sigma(t)$  are piecewise constant, we only need to specify their values  $\{a_i\}_{i=0..N-1}$  and  $\{\sigma_i\}_{i=0..N-1}$  at the points  $\{t_i\}_{i=0..N-1}$ , with the convention on  $[0, +\infty)$

$$a(t) = a_i \quad t \in [t_i, t_{i+1}), \quad i = 0..N-2 \quad (20)$$

$$a(t) = a_{N-1} \quad t \geq t_{N-1} \quad (21)$$

$$\sigma(t) = \sigma_i \quad t \in [t_i, t_{i+1}), \quad i = 0..N-2 \quad (22)$$

$$\sigma(t) = \sigma_{N-1} \quad t \geq t_{N-1}. \quad (23)$$

Let us illustrate this by assuming the time grid runs up to 30 years with 6M intervals. There are then  $N = 61$  points in the grid ( $t_0 = 0, t_{60} = 30$ ). A generic piecewise constant function would thus be defined by 61 values, which means 122 parameters including the mean reversion and the volatility. This is clearly too much, so one must constrain them in some manner.

One natural way to reduce the number of parameters is to match their number with the number of market maturities, with the view to calibrating using a bootstrap-like method. If the first market maturity is at  $t_2 = 1$ , then one will enforce the conditions  $a_1 = a_0$  and  $\sigma_1 = \sigma_0$ , and so on and so forth for the remaining maturities. Although this strategy is commonly used, probably due to its intuitive meaning and ease of implementation, we are not interested in developing it further here. It suffers from serious drawbacks, among which big jumps in values from one time to the other, causing instabilities, especially visible when pricing exotics.

We choose instead to constrain the parameters  $a_i$  and  $\sigma_i$  to be generated by functional forms. The resulting number of free parameters will be that of the functional forms, while a suitable choice of such functional form will ensure that the functions  $a(t)$  and  $\sigma(t)$  do not suffer from dangerously jumping values.

Apart from the constant case, in this work we will make use of the logistic function to generate the mean reversion, i.e. we define

$$a_i = A_0 + \frac{A_1 - A_0}{1 + e^{A_2(A_3 - t_i)}} \quad (24)$$

where  $A_0..A_3$  are 4 parameters describing (heuristically)

- $A_0$ : the short-term value
- $A_1$ : the mid/long term value

- $A_2$ : the slope at the transition
- $A_3$ : the time at the transition.

We call transition the area where the function changes rapidly between its early time and late time regimes. See fig. 1 for a visual example of the type of shapes it produces.

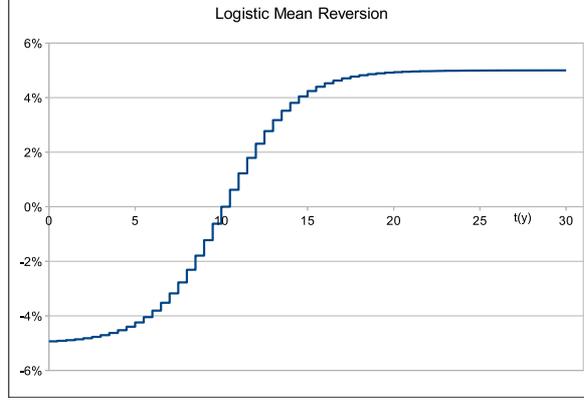


Figure 1: Piecewise constant mean reversion based on the logistic function  $\text{Logistic}[-5\%, 5\%, 0.5, 10]$

For the volatility we will study a functional form based on interpolating at the times  $\{t_i\}_{i=0..N-1}$  a given grid of volatility vs. time  $\{\tilde{T}_k, \Sigma_k\}_{k=0..\tilde{N}-1}$ , i.e.

$$\sigma_i = \text{CubicSpline}(t_i, \{\tilde{T}_k\}, \{\Sigma_k\}) \quad (25)$$

where we interpolate with the Cubic Spline with left constraint of vanishing 2nd derivative and right constraint of vanishing 1st derivative. The volatility  $\sigma(t)$  is set to constant after the last time  $\tilde{T}_{\tilde{N}-1}$ .

### 3 Calibration Strategies

In this section we describe a few possible strategies for the calibration of Hull-White model to the market of swaptions. By calibration strategy we mean the following points:

1. The choice of constant or time-dependent mean reversion and volatility
2. The choice of products to calibrate to, and whether to calibrate locally or globally
3. Whether to optimize on the mean reversion  $a(t)$  and the volatility  $\sigma(t)$  together or separately, and in the latter case, how to estimate one independently of the other.

#### 3.1 Relations Market/Model parameters

Before diving into the calibration procedures, it is useful to investigate what types of swaption implied volatilities can be obtained from Hull-White model. Indeed the range of accessible volatilities is restricted, as it is not a market model such as LMM or SMM for which the market implied volatilities are inputs. On the contrary Hull-White model constrains the shape of the implied volatility matrix, and our goal in this section is to study the influence of the mean reversion and the volatility on this matrix.

First of all, we take a fixed constant mean reversion (1%) and change the value of the constant volatility  $\sigma$ . We display the result in fig. 2, for the two maturities 6M and 10Y.

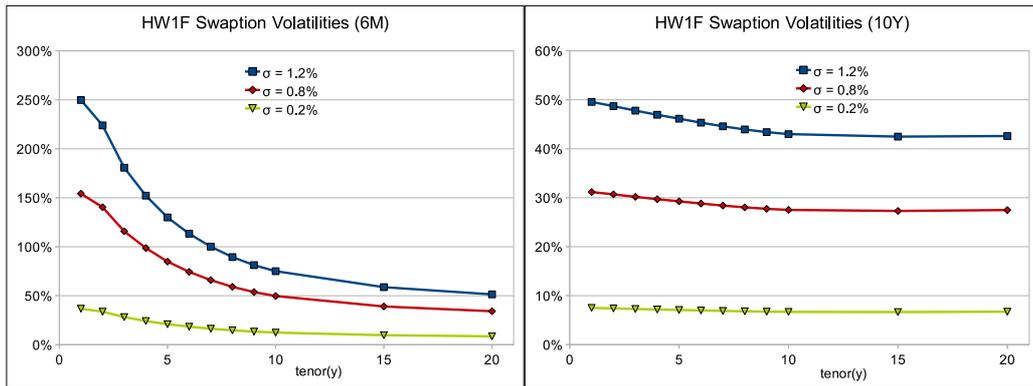


Figure 2: Implied Volatilities for model swaptions with different volatilities. Maturities 6M and 10Y, mean reversion  $a = 1\%$ .

Apart from the intuitive fact that the implied volatility increases when the model volatility increases, we can observe that, at fixed maturity, the ratios between two implied volatilities with neighbouring tenors do not vary much with  $\sigma$ . This is an encouraging fact when considering the use of SMM approximation and will be discussed in more details in section 4.1. It means that  $\sigma$  influences mostly the level of the implied volatility curves, without changing their shapes much.

Next we fix  $\sigma$  and vary the mean reversion. We find the results in fig. 3, which suggest that the mean reversion has a qualitatively different influence on the implied volatility. Indeed, this time the shape is modified. Considering the wide change between  $a = 5\%$  and  $a = -4\%$  we can see that even the monotonicity of the implied volatility curve can change due to the mean reversion.

Note also that the level of the implied volatility decreases significantly when the mean reversion  $a$  increases.

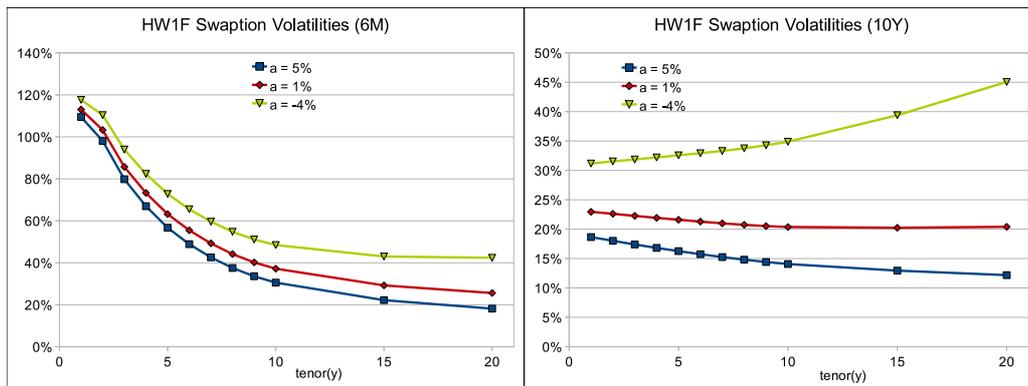


Figure 3: Implied Volatilities for model swaptions with different mean reversions. Maturities 6M and 10Y, volatility  $\sigma = 0.6\%$ .

Remark that until now the implied volatility curves have been monotonic. Considering that the market may have a hump, we would like to know whether such a shape is accessible in Hull-White model, i.e. whether it is possible to see a change in monotonicity. A quick look at fig. 3 shows us that for positive mean reversions, the implied volatility is decreasing with the tenor, but for negative mean reversions, the volatility can be increasing. It is then natural to expect that a hump could appear if a time-dependent mean reversion started in low (possibly negative) values and ended in higher values. This is actually the main motivation for introducing the logistic parametric form in section 2.4. In fig. 4 we show a few examples of hump and the corresponding model parameters.

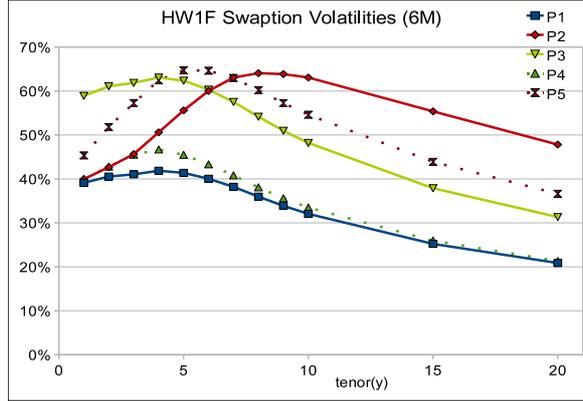


Figure 4: Implied Volatilities for model swaptions with different time-dependent mean reversions. Maturity 6M, model parameters

P1:  $\sigma = 0.2\%$   $a = \text{Logistic}[-30\%, 5\%, 1, 3]$ , P2:  $\sigma = 0.2\%$   $a = \text{Logistic}[-30\%, 5\%, 1, 6]$   
P3:  $\sigma = 0.3\%$   $a = \text{Logistic}[-30\%, 5\%, 1, 3]$ , P4:  $\sigma = 0.2\%$   $a = \text{Logistic}[-30\%, 5\%, 10, 3]$   
P5:  $\sigma = 0.2\%$   $a = \text{Logistic}[-50\%, 5\%, 1, 3]$ , P6:  $\sigma = 0.1\%$   $a = \text{Logistic}[-50\%, 5\%, 1, 3]$

All curves in fig. 4 are obtained from one another by changing the value of one of the parameters  $\sigma$  and  $A_i$ ,  $i = 0..3$ . Comparing curves P1 and P2 shows that we can control the position of the hump in time with parameter  $A_3$ . Note however that since this makes the mean reversion negative on a longer range, the overall level of curve P2 has increased. This effect may be compensated by a decrease in  $\sigma$ .

Curves P1 and P3 provide an other example that increasing only the volatility gives a proportional level change.

Curves P1 and P4 show that one can control, to some limited extent though, the steepness of the hump by changing the parameter  $A_2$ .

Curves P1, P5 and P6 show how to obtain a hump with bigger amplitude by taking more negative values for the mean reversion at early times (P5) and reduce the implied volatility level by decreasing the volatility.

From these examples we can see that a rather wide range of humped shapes can be obtained by suitably setting the values of the parameters of the logistic function  $A_i$ ,  $i = 0..3$ , while compensating with the volatility  $\sigma$  to adjust the level.

Overall this tells us that the mean reversion and volatility must be considered as a set, each of them having a different impact on the implied volatilities. Moreover, both of them influence the level of these volatilities, a fact that should be kept in mind when trying to build an intuition of what values these parameters can take and how they vary in day-to-day market changes. For instance, the same volatility  $\sigma$  can be considered as big or small, depending on the value of the corresponding mean reversion.

### 3.2 Optimizing on the Mean Reversion or not

To our knowledge, there are two methods commonly adopted by market practitioners for calibrating the HW1F model to the vanillas. The first one involves fixing the mean reversion using some correlation input and then calibrating only the volatility parameter to cap/floorlets and/or swaption prices. This strategy is mainly motivated by the theoretical observation that the mean reversion

has a strong influence on rate correlations, see eq. (46), and therefore on exotics such as Bermudan swaptions. This was demonstrated in [4], and we also briefly address this last issue in section 4.7.

On the other hand, the second method attempts to calibrate both the mean reversion and the volatility parameters to a series of cap/floorlets and/or swaptions with different maturities and tenors. This is based on the fact that both model parameters have crucial impacts on cap/floorlets and swaption prices, as is evident from the numerical examples given in the previous section. An intuitive way to see this is that both the mean reversion and the model volatility parameterize the associated Heath–Jarrow–Morton (HJM) volatility function.

In the current work, we focus mainly on the second calibration method. However, we shall emphasize that the objective of this study is not to reach a conclusion as to which strategy is superior to the other. Our intention is rather to lay out the methodology and gather some concrete numerical results, which we believe is valuable from a practical point of view as such results are scarce in the existing literature.

For the sake of completeness and in order to have a frame of reference, we have also included in our analysis some numerical examples of the first calibration method described above<sup>1</sup>.

### 3.3 Bootstrap vs. Overall Calibration

The first strategy that comes to mind is of bootstrap type, i.e. to use the model parameters  $a_i$  and  $\sigma_i$  from 0 to the first market maturity  $T_1$  to fit to the first instrument, and then the model parameters from  $T_1$  to  $T_2$  to fit the second instrument at maturity  $T_2$ , and so on and so forth.

Although the above method may feel natural, produce perfect fit for (at least some of) the chosen instruments, and has the advantage of requiring only low-dimension optimizations, it often results in big jumps in the parameters and tends to exhibit a high dependency on the initial values fed to the optimization algorithm. We consider these issues as serious drawbacks and prefer adopting the following strategy.

We give up the intention of exact-fitting some specific instruments, and consider them instead as a pool on which to optimize. We also choose parametric forms for the mean reversion and volatility, as explained in the previous section, with a number of model parameters and fitted instruments not necessarily equal. We then optimize on all model parameters to fit all the chosen market instruments at once, with a predefined distance (here  $\mathcal{L}_2$ ) between the model and market prices, together with weights to focus on some instruments more than others.

The parametric forms actually put constraints on all parameters as a whole, such that if we choose the functionals suitably, no big jumps will appear. The drawback of this is that there is no more direct relation between parameters and instruments, and in general we need to perform multi-dimensional optimizations, which are more difficult to handle than one-dimensional algorithms. It is one of the purposes of this document to prove that these methods are feasible and can produce more stable and meaningful results.

### 3.4 Calibration to Swaptions

Here we present three ways of proceeding, independently of whether the parameters are constant or time-dependent. The main difference between these methods is the treatment of the mean reversion. In the first method, we fix it. In the second method, we estimate it independently of the volatility, and then find the volatility. In the third method, we optimize on the mean reversion and the volatility simultaneously. All multi-dimensional optimizations are performed using a version of

---

<sup>1</sup>For our purposes, it is sufficient to fix the mean reversion by hand without using correlation input.

the Simulated Annealing algorithm, see appendix B. Below are more detailed descriptions of these three calibration methods.

### 3.4.1 Method 1

In this calibration method we do not optimize on the mean reversion, which is fixed to some arbitrary (constant) value to be specified by the user. We therefore optimize on the volatility only, such that the analytical price  $PV_{\mathbf{a}}^{\text{mod}}(M_i, T_j)(\sigma)$  approaches the market price  $PV^{\text{mkt}}(M_i, T_j)$ .  $PV^{\text{mod}}$  is the price of a swaption as calculated in eq. (12), at ATM, and the subscript  $\mathbf{a}$  indicates that the mean reversion has been fixed previously. In other words, our calibrated volatility is the point  $\sigma$  (possibly multi-dimensional) at which the function

$$G_{\mathbf{a}}(\sigma) = \sum_{1 \leq i \leq n_m, 1 \leq j \leq n_t - 1} W_{i,j} \left( \frac{PV_{\mathbf{a}}^{\text{mod}}(M_i, T_j)(\sigma)}{PV^{\text{mkt}}(M_i, T_j)} - 1 \right)^2 \quad (26)$$

has its minimum. The initial point of the optimization must be chosen by the user. For the time-dependent case, one can first optimize for a constant  $\sigma$  and use the result as a hint for the initial point of the multi-dimensional optimization, in order to improve the speed of convergence.  $W_{i,j}$  are the user-defined weights.

### 3.4.2 Method 2

The second method is divided into two steps:

1. Estimate the mean reversion  $a(t)$  using SMM approximation for the implied volatilities
2. Optimize on the volatility using the exact analytical price, knowing the mean reversion from step 1).

Step 1) above is rendered possible, without knowledge of the volatility, thanks to property (8) of SMM approximated implied volatilities. Indeed, taking the ratio of two implied variances with the same maturity  $M_i$  but different tenors  $T_j$  and  $T_k$ , we find,

$$\frac{V_{\text{swap}}(M_i, T_j)}{V_{\text{swap}}(M_i, T_k)} = \left[ \frac{(P(0, M_i) - P(0, T_k))B(M_i, T_j)}{(P(0, M_i) - P(0, T_j))B(M_i, T_k)} \right]^2 \quad (27)$$

which is independent of the volatility  $\sigma(t)$ , as can be seen in the definition of  $B(t, T)$  in (31).

This means that by optimizing such that ratios of the type (27) approach their market counterpart, one can obtain an estimation of the mean reversion without knowing the volatility function  $\sigma(t)$ . More precisely, let us denote by  $IV_{i,j}$  the market implied volatility for the swaption of maturity  $M_i$  and tenor  $T_j$ , with  $n_m$  maturities and  $n_t$  tenors. Our calibrated mean reversion is the point  $\mathbf{a}$  (possibly multi-dimensional) at which the function

$$F(\mathbf{a}) = \sum_{1 \leq i \leq n_m, 1 \leq j \leq n_t - 1} \frac{W_{i,j+1}}{W_{i,j}} \left( \sqrt{\frac{V_{\text{swap}}(M_i, T_{j+1})}{V_{\text{swap}}(M_i, T_j)}}(\mathbf{a}) - \frac{IV_{i,j+1}}{IV_{i,j}} \right)^2 \quad (28)$$

has its minimum. In case a weight  $W_{i,j} = 0$  is met, it is simply ignored in the sum as we do not intend to use the corresponding swaption for calibration. The index  $j$  will then be chosen as the next one such that  $W_{i,j} \neq 0$ .

Step 2) is an optimization on the volatility  $\sigma(t)$  such that the analytical price approaches its market counterpart. In other words, our calibrated volatility is the point  $\sigma$  (possibly multi-dimensional) at which the function defined in (26) has its minimum, with the mean reversion fixed at step 1).

### 3.4.3 Method 3

The third method is conceptually more simple and goes into one step only: optimize on the mean reversion *and* the volatility such that the analytical price approaches the market price. In other words, our calibrated mean reversion and volatility are the points  $\mathbf{a}$  and  $\sigma$  (possibly multi-dimensional) at which the function

$$G(\mathbf{a}, \sigma) = \sum_{1 \leq i \leq n_m, 1 \leq j \leq n_t - 1} W_{i,j} \left( \frac{\text{PV}^{\text{mod}}(M_i, T_j)(\mathbf{a}, \sigma)}{\text{PV}^{\text{mkt}}(M_i, T_j)} - 1 \right)^2 \quad (29)$$

has its minimum. To choose the initial point of the optimization algorithm (multi-dimensional in any case), we have two options. We can either choose it based on intuition and experiment, similarly to the previous methods, or first estimate a mean reversion and volatility with method 2 and use the results as initial guesses. This guess usually yields a clear improvement in the speed of convergence.

## 4 Numerical Results

In this section we display numerical results of calibrations to JPY ATM swaptions using the various methods described in Section 3. Similar strategies can be applied to other currencies. We show examples of calibration to USD and EUR swaption markets in Appendix C.3 and C.4.

### 4.1 Performance of SMM approximation

The SMM approximation formula (19) has shed some light on the roles of the mean reversion and volatility parameters in determining implied swaption volatilities. In this section, we attempt to verify the implication of the SMM approximation formula and its accuracy through some numerical examples.

In the left panel of Figure 5 below, we display the ratios of the implied volatilities of swaptions with a 5-year maturity and adjacent tenors under various mean reversion parameters. Here, we fix the volatility parameter to 0.5%. The right panel focuses on the impacts of changing the volatility parameters on the implied volatility ratios with the mean reversion fixed to 8%. All implied volatilities used in Figure 5 are computed using the Jamshidian decomposition method. It is clear that the implied volatility ratio curve moves systematically downward as the mean reversion parameter increases. In contrast, the changes in implied volatility ratios are relatively insignificant as the volatility parameter varies. This observation indeed agrees with what the SMM approximation formula suggests: implied volatility ratios of swaptions with the same maturity date depend mainly on the mean reversion parameter. It is this observation that motivates the introduction of the second calibration strategy described in the previous section in which the mean reversion and volatility parameters can be calibrated separately.

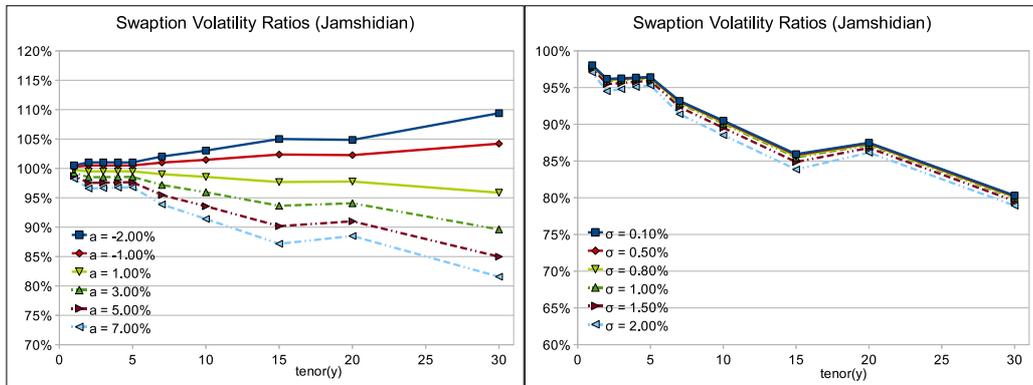


Figure 5: Swaption Implied Volatility Ratios, maturity 5Y.

Left: changing  $a$  at fixed  $\sigma$

Right: changing  $\sigma$  at fixed  $a$ .

In Figures 23 in Appendix C.1, we also compare the swaption implied volatilities computed using the Jamshidian method with those computed using the SMM approximation formula. The top two panels correspond to two different volatility parameter values (0.5% and 1.5%), while fixing the mean reversion at 8%. On the other hand, the bottom two panels correspond to two different mean reversion values (-2% and 3%), while fixing the volatility at 0.5%. In all four figures, we fix

the maturity of the swaptions to 5 years while changing the underlying swap tenors from 1 year to 30 years. From these figures, we see that the SMM approximations appear to be able to capture the qualitative behavior of the benchmark swaption implied volatilities obtained from the Jamshidian method. However, under certain cases, the approximated swaption implied volatilities are quite far from the benchmark values calculated from the Jamshidian method. For example, in the third panel, the discrepancies between the SMM approximated volatilities and the benchmark values can be more than 10% for swaptions with long tenors. Indeed, we have run the same comparison tests under many other parameter values and the results agree what we have suggested before: the SMM approximation can be crude and is not meant to be used for actual pricing purposes.

We emphasize that the SMM approximation shall be used as a means to understand the qualitative behavior of the implied swaption volatilities in terms of the HW model parameters, which can be useful from model control point of view. Its accuracy can vary very much depending on the model parameter values and the tenor and maturity of the underlying swaption. Hence, it cannot serve as a reliable way for pricing swaptions in general. In this next section, we shall also demonstrate numerically the effectiveness of the calibration strategy utilizing the SMM approximation result.

## 4.2 Choice of Instruments

As an illustration consider the case of a calibration to JPY swaptions, for which we have 10 maturities and 12 tenors, in other words 120 swaptions. That is not to say there are 120 degrees of freedom (d.o.f.). As it is not clear what the actual number of d.o.f. is, we do not know exactly how many parameters in the model can be used to fit how many swaptions. This requires testing.

In the tests we present here, we will consider 3 instrument choice scenarios. The first one represents a strategy commonly used, the local calibration, in parts for hedging reasons, which consists in picking swaptions with fixed co-terminal (maturity + tenor). We will consider here the 20Y co-terminal calibration. In practice this means that we choose non-zero weights  $W_{i,j}$  only for one swaption per maturity.

The second scenario is an extension of the previous one where we consider 2 co-terminal dates, 10Y and 20Y. The purpose of this test is to analyse the behaviour of the calibration procedure when we increase the quantity of information and see how this translates in terms of model parameters. This is rather a toy strategy, we are not aware of its actual use in the industry.

Finally the third scenario represents the calibration to the whole swaption matrix, i.e. global calibration. The goal here is to take into account all the market information and analyse, on the one hand how many model parameters are needed to describe it and on the other hand what is the best fitting quality that we can hope for, knowing that perfect fit to the whole matrix cannot be expected in this model.

We will consider a testing period of 2 years between Nov. 2007 and Sept. 2009 where we will perform the calibration weekly.

## 4.3 Co-terminal 20Y

According to our interpretation of the mean reversion as influencing strongly the slope of the swaption implied volatility curves, in the tenor direction, it seems risky to calibrate it in the settings described here. Indeed, the only information we have is 1 number for each maturity, such that we have no information at all on the slope.

Taking the simplest case of constant  $a$  and  $\sigma$ , we do find a typical behaviour of over-fitting on the mean reversion side. In fig. 24 in the appendix, we show the constant mean reversion calibrated

over our testing period. It bears big jumps from weeks to weeks such that even its overall order of magnitude is unclear. Although it is not obvious what this order of magnitude should be, such jumps do not seem acceptable for pricing in practice. We will therefore conclude that it is too unstable to optimize on the mean reversion when only 1 instrument per maturity is chosen, such that in the remainder of this subsection, we keep the mean reversion fixed over the testing period.

Now we fix the mean reversion at the value obtained on Nov. 27th, 2007 with Method 3, i.e.  $a = 0.97\%$ , and proceed to calibrating the volatility as a constant or time-dependent function. We are interested in 2 aspects, the stability of the parameters over the 2 year period of our tests, and the quality of the fit to the considered market instruments. For the volatility function  $\sigma(t)$  we first consider a 10pts function with no constraints on the parameters, such that, having the same number of parameters as fitted instruments, we can hope for a perfect fit. This is indeed what happens, as can be seen in table 1 below, which gives an example of fit quality comparison between the constant and time-dependent calibrations. "Mkt" refers to the market implied volatilities, "Md" to the model implied volatilities for constant  $\sigma$ , and "Md<sub>t</sub>" to the model implied volatilities for time-dependent  $\sigma(t)$ .

Table 1: 20Y Fitted JPY swaptions on 2009/09/08, no parameter constraints

Inst.	Mkt	Md	Md <sub>t</sub>	Md-Mkt	Md <sub>t</sub> -Mkt
1M20Y	29.00%	28.92%	29.25%	-0.08%	0.25%
3M20Y	29.30%	28.67%	29.03%	-0.63%	-0.27%
6M20Y	29.40%	28.33%	29.41%	-1.07%	0.01%
1Y20Y	28.80%	27.59%	28.80%	-1.21%	0.00%
2Y20Y	27.10%	26.28%	27.10%	-0.82%	0.00%
3Y15Y	26.20%	27.01%	26.20%	0.81%	0.00%
4Y15Y	25.00%	25.60%	25.00%	0.60%	0.00%
5Y15Y	24.00%	24.50%	24.00%	0.50%	0.00%
7Y15Y	22.70%	23.10%	22.70%	0.40%	0.00%
10Y10Y	21.90%	22.66%	21.90%	0.76%	0.00%

Apart from the earliest maturities, which are not very liquid anyway, the fit is nearly perfect in the case of the time-dependent model volatility. However a study of the stability of the parameters over the testing period shows in fig. 25 in the appendix that, while the constant volatility  $\sigma$  and the parameter  $\Sigma_0$  of  $\sigma(t)$  are stable, the parameter  $\Sigma_9$  of  $\sigma(t)$  is very noisy. We conclude that although we can reach a nearly perfect fit for most chosen instruments, there is again over-fitting, this time due to  $\sigma(t)$ .

There are two ways to remedy this problem, one is reducing the number of parameters in our parametric form for  $\sigma(t)$ , and the other one is putting constraints on these. Since both strategies lead to similar behaviours, we choose to display only the latter here. We choose to constrain the parameters  $\Sigma_i$  such that  $|\Sigma_{i+1} - \Sigma_i| < \alpha \Sigma_i$  for  $\alpha = 0.1$ . We obtain the stability results in fig. 6

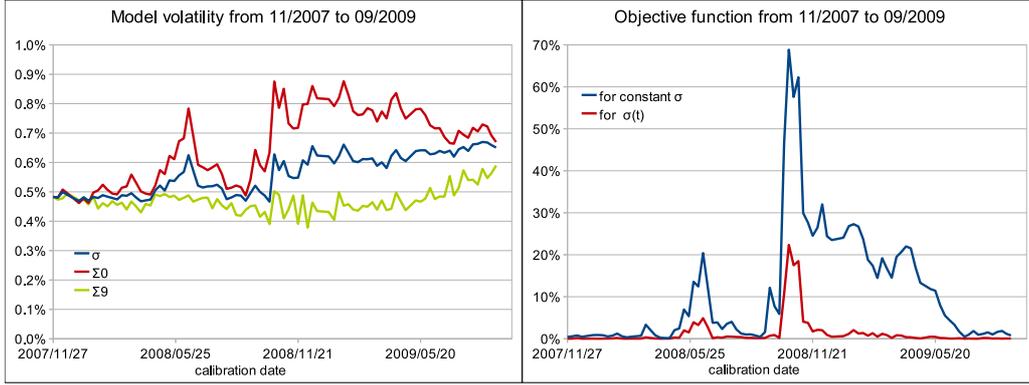


Figure 6: Volatility calibrated to 20Y co-terminal at fixed mean reversion. Constraint  $\alpha = 0.1$ .  
 Left: constant  $\sigma$ , parameters  $\Sigma_0$  and  $\Sigma_9$  of  $\sigma(t)$   
 Right: objective function for constant  $\sigma$  and time-dependent  $\sigma(t)$ .

Once the constraint has been enforced, the calibration result becomes stable. As should be expected, the time-dependence of the volatility always improves the quality of the fit, though this improvement may be more or less significant depending on the dates. It appears to be particularly efficient in times of unstable or quickly changing markets, for instance at and after Lehman crisis at the fall of 2008.

In terms of quality of fitting to the chosen instruments, we display in table 2 the result of the calibration on the dates 2009/04/14 and 2009/09/08. As can be seen in fig. 6, 2009/04/14 is typical of the "after-Lehman-shock" period, extending over several months, during which calibration with constant parameters behaves relatively poorly.

Table 2: 20Y Fitted JPY swaptions in 2009,  $\alpha = 0.1$

Inst.	Mkt 04/14	Md-Mkt	Md <sub>t</sub> -Mkt	Mkt 09/08	Md-Mkt	Md <sub>t</sub> -Mkt
1M20Y	38.20%	-9.21%	-1.07%	29.00%	-0.08%	0.25%
3M20Y	38.10%	-9.35%	-1.33%	29.30%	-0.63%	-0.27%
6M20Y	34.10%	-5.67%	0.49%	29.40%	-1.07%	0.04%
1Y20Y	30.20%	-2.45%	0.75%	28.80%	-1.21%	0.02%
2Y20Y	27.00%	-0.46%	-0.06%	27.10%	-0.82%	-0.31%
3Y15Y	26.00%	1.71%	0.45%	26.20%	0.81%	0.35%
4Y15Y	24.30%	2.09%	0.02%	25.00%	0.60%	-0.12%
5Y15Y	23.00%	2.31%	0.06%	24.00%	0.50%	0.04%
7Y15Y	21.10%	2.72%	0.18%	22.70%	0.40%	-0.03%
10Y10Y	19.90%	3.39%	0.22%	21.90%	0.76%	0.01%

On the 2009/04/14 the time-dependent calibration offers a clear improvement over the constant one, although now, due to the presence of the constraint on the cubic spline, perfect fit is no longer possible. This improvement will not always be so obvious, as can be observed on the later date 2009/09/08.

### 4.3.1 Conclusion

We conclude from these tests that the time-dependency of the volatility enables us to adapt in a more flexible way to different types of swaption matrices that may appear in times of large market moves. Some quality of fitting has been lost compared to the unconstrained volatility case, but stability has greatly improved.

Although we have chosen a specific set of instruments and analysed the fitting quality for these instruments, it is interesting to have a look at how the other swaptions were fit. In fig. 7, we display the market and model swaptions at the 1Y and 10Y maturities, for the time-dependent volatility with constraint calibration scheme. We can see that away from the chosen instruments, the model swaptions can be off by a very large amount, especially for early maturities and in the presence of a hump.

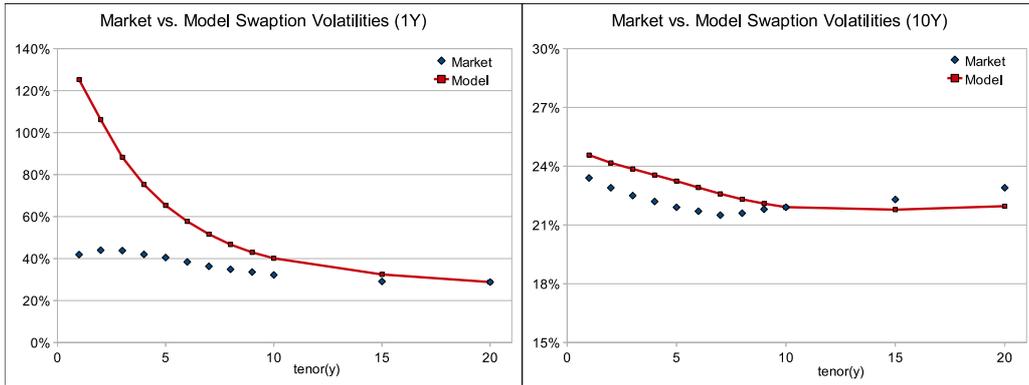


Figure 7: Market and Model Implied Volatilities on 2009/09/08, constant  $a$ , time-dependent  $\sigma(t)$   
 Left: maturity 1Y, calibrated instrument at tenor 20Y  
 Right: maturity 10Y, calibrated instrument at tenor 10Y

## 4.4 Co-terminals 10Y and 20Y

### 4.4.1 Constant $a$ and $\sigma$

Here we compare the results of the 3 methods described in section 3.4. We briefly summarize them below.

- **Method 1** The mean reversion is fixed and the volatility is calibrated by using the analytical prices. This means we perform one 1-dimensional minimization.
- **Method 2** The mean reversion is calibrated separately by using the approximated implied volatility ratios, while the volatility is calibrated by using the analytical prices. This means we perform two 1-dimensional minimizations successively.
- **Method 3** The mean reversion and volatility are calibrated simultaneously by using the analytical prices. This means we perform one 2-dimensional minimization.

First we display the 3 different mean reversions and model volatilities obtained with these methods in fig. 8

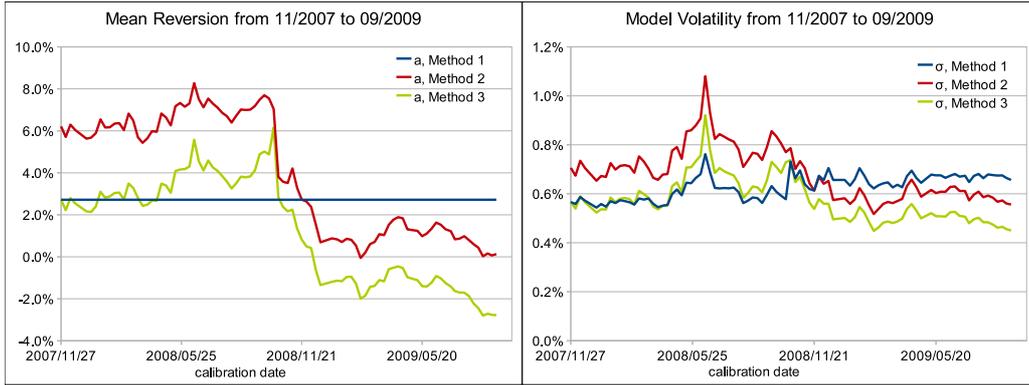


Figure 8: Constant  $a$  and  $\sigma$  are calibrated to the co-terminals 10Y and 20Y.

We observe that all these methods have a similar satisfactory stability. On top of this, we see a clear regime change in the mean reversion at the time of Lehman crisis. After this, the mean reversion is clearly lower than before. This echoes a structural change in the swaption matrix that will be discussed in more details in section 4.6.

As to the quality of the fit to the chosen swaptions, we find the target values in fig. 9 for all dates and the detailed fitting errors in table 3 for the particular date 2009/09/08.

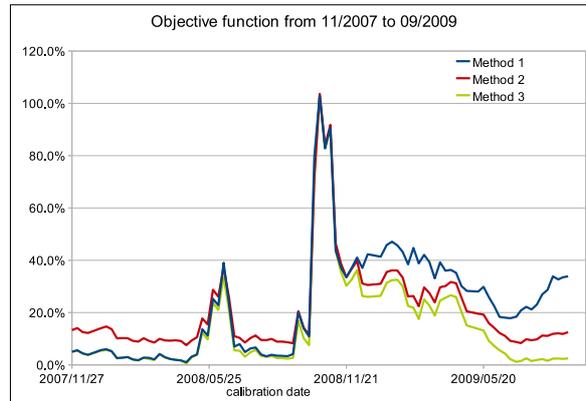


Figure 9: Constant  $a$  and  $\sigma$  are calibrated to the co-terminals 10Y and 20Y.

Table 3: 10Y and 20Y co-terminal fitted JPY swaptions on 2009/09/08

Inst.	Mkt	Md(1)-Mkt	Md(2)-Mkt	Md(3)-Mkt
1M10Y	33.40%	6.52%	4.89%	2.48%
1M20Y	29.00%	-3.97%	-2.31%	-0.28%
3M10Y	34.00%	5.04%	3.52%	1.23%
3M20Y	29.30%	-4.52%	-2.82%	-0.75%
6M10Y	33.60%	4.28%	2.92%	0.80%
6M20Y	29.40%	-4.96%	-3.21%	-1.05%
1Y9Y	33.60%	4.53%	2.95%	0.58%
1Y20Y	28.80%	-5.10%	-3.25%	-0.95%
2Y8Y	31.90%	3.58%	2.13%	-0.09%
2Y20Y	27.10%	-4.72%	-2.67%	-0.10%
3Y7Y	29.60%	3.41%	2.07%	0.00%
3Y15Y	26.20%	-2.64%	-1.44%	-0.27%
4Y6Y	28.10%	2.82%	1.56%	-0.37%
4Y15Y	25.00%	-2.86%	-1.45%	0.04%
5Y5Y	26.70%	2.37%	1.19%	-0.62%
5Y15Y	24.00%	-2.98%	-1.37%	0.42%
7Y3Y	24.80%	1.25%	0.18%	-1.40%
7Y15Y	22.70%	-3.21%	-1.20%	1.24%
10Y1Y	23.40%	-0.08%	-0.79%	-1.83%
10Y10Y	21.90%	-2.58%	-0.96%	0.91%

Several conclusions can be drawn from these numbers. Up to the Lehman crisis time, Method 1 and Method 3 behave quite similarly. This is natural as the fixed mean reversion is the calibrated one on the first day, and the mean reversion with Method 3 does not vary much during this period, such that its original value on 2007/11/27 remains a good estimation. However from the time the crisis is triggered, we have seen that the mean reversion enters a new regime. Method 1 is then no longer able to adapt to the new market conditions. Method 2 on the other hand, while fitting not as well as Method 3 since it is not a true 2-dimensional optimization, is nevertheless able to take into account the new situation after the shock.

We conclude that the mean reversion can be estimated safely when two co-terminal swaption information is included, and that it should be optimized on if one wants to adapt to changing market conditions. While method 3 is a true 2-dimensional optimization and therefore offers the best fitting quality in these parameter settings, it is clearly slower at runtime. Method 2 represents an alternative for which we can both adapt to market changes and have a very short runtime (quasi-instantaneous), if one is ready to sacrifice a little fitting quality.

#### 4.4.2 Constant $a$ and Time-Dependent $\sigma(t)$

Similarly to the constant  $\sigma$  case, we can distinguish the 3 methods

- **Method 1:** the mean reversion is fixed, and we perform one multi-dimensional optimization on the parameters  $\Sigma_k$  of  $\sigma(t)$
- **Method 2:** the mean reversion is calibrated separately by a 1-dimensional optimization, and we perform one multi-dimensional optimization on the parameters  $\Sigma_k$  of  $\sigma(t)$

- **Method 3:** the mean reversion and the volatility are calibrated simultaneously in one multi-dimensional optimization on  $a$  and the parameters  $\Sigma_k$  of  $\sigma(t)$ .

First of all, let us compare the constant and time-dependent volatility calibrations. In fig. 10 we display the targets over the test period. In order not to burden the text with too many graphics, we show only the result for Method 3, similar conclusions holding for the other methods.

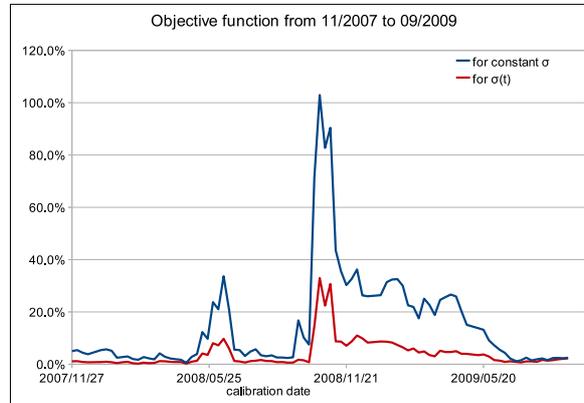


Figure 10: Calibration to co-terminals 10Y and 20Y, constant  $a$ , Method 3.

As can be expected, time-dependence represents a clear improvement in general, though for some special dates it might be less significant. Indeed, it may happen that the global minimum of the target function is obtained for an almost constant  $\sigma(t)$ . Such situations occur in the summer of 2009, as can be seen in the figure above. Outside of these periods, the improvement due to time-dependence is obvious. We display the detailed fitting results for such a case in table 4.

Table 4: 10Y and 20Y fitted JPY swaptions on 2009/04/14,  $\alpha = 0.1$

Inst.	Mkt	Md-Mkt	Md <sub>t</sub> -Mkt
1M10Y	42.20%	-3.65%	3.34%
1M20Y	38.20%	-9.40%	-3.03%
3M10Y	43.20%	-5.23%	1.62%
3M20Y	38.10%	-9.51%	-3.22%
6M10Y	40.40%	-3.18%	1.43%
6M20Y	34.10%	-5.78%	-1.19%
1Y9Y	36.80%	0.34%	1.48%
1Y20Y	30.20%	-2.46%	-0.54%
2Y8Y	33.80%	1.34%	0.06%
2Y20Y	27.00%	-0.29%	-0.19%
3Y7Y	31.80%	1.64%	-0.28%
3Y15Y	26.00%	1.23%	0.40%
4Y6Y	30.40%	1.60%	-0.91%
4Y15Y	24.30%	1.81%	0.55%
5Y5Y	29.20%	1.52%	-1.20%

Inst.	Mkt	Md-Mkt	Md <sub>t</sub> -Mkt
5Y15Y	23.00%	2.22%	0.84%
7Y3Y	27.20%	1.22%	-1.90%
7Y15Y	21.10%	2.97%	1.29%
10Y1Y	25.10%	0.74%	-2.17%
10Y10Y	19.90%	3.39%	1.47%

Note also that now that we fit to a larger number of swaptions, a near-perfect fit can no longer be hoped for, considering the number of model parameters available here.

In fig. 11 we compare the fit to the co-terminals 10Y and 20Y for the 3 calibration methods described in this subsection, all with a time-dependent model volatility.

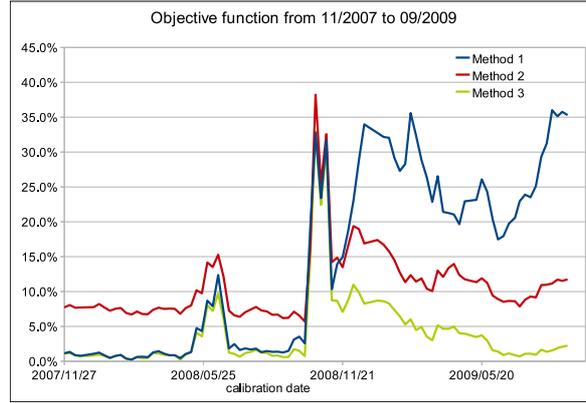


Figure 11: Calibration to co-terminals 10Y and 20Y, constant  $a$  and time-dependent  $\sigma(t)$

We find that these targets reproduce a similar pattern as for the constant volatility case. Consequently the same conclusions hold, i.e. that calibrating the mean reversion is preferable, especially in case a wildly changing markets, and that simultaneous calibration performs better in terms of fitting quality than separate calibration of the mean reversion and volatility. For the particular date 2009/09/08, the detailed errors for each method are given in table 5 below.

Table 5: 10Y and 20Y co-terminal fitted JPY swaptions on 2009/09/08

Inst.	Mkt	Md(1)-Mkt	Md(2)-Mkt	Md(3)-Mkt
1M10Y	33.40%	5.14%	3.75%	1.76%
1M20Y	29.00%	-5.11%	-3.11%	-0.89%
3M10Y	34.00%	3.73%	2.44%	0.55%
3M20Y	29.30%	-5.62%	-3.58%	-1.33%
6M10Y	33.60%	4.38%	3.00%	1.02%
6M20Y	29.40%	-5.18%	-3.15%	-0.91%
1Y9Y	33.60%	4.55%	3.05%	0.85%
1Y20Y	28.80%	-5.40%	-3.18%	-0.77%
2Y8Y	31.90%	4.04%	2.45%	0.03%
2Y20Y	27.10%	-4.76%	-2.44%	-0.03%

Inst.	Mkt	Md(1)-Mkt	Md(2)-Mkt	Md(3)-Mkt
3Y7Y	29.60%	3.03%	1.72%	0.17%
3Y15Y	26.20%	-3.14%	-1.71%	-0.15%
4Y6Y	28.10%	2.77%	1.51%	-0.19%
4Y15Y	25.00%	-3.14%	-1.49%	0.17%
5Y5Y	26.70%	2.74%	1.30%	-0.54%
5Y15Y	24.00%	-3.00%	-1.28%	0.46%
7Y3Y	24.80%	2.40%	0.74%	-1.42%
7Y15Y	22.70%	-2.69%	-0.72%	1.18%
10Y1Y	23.40%	1.49%	0.13%	-1.49%
10Y10Y	21.90%	-1.57%	-0.12%	1.25%

In terms of runtime, these 3 methods differ only by their treatment of the mean reversion. Since the volatility is time-dependent, in any case we must run a multi-dimensional constrained optimization, the question is whether to run it on 10 parameters (volatility only) or 11 (volatility and mean reversion). The runtime difference is not such that one can consider Method 2 as significantly faster than Method 3.

#### 4.4.3 Time-Dependent $a(t)$ and $\sigma(t)$

For both available methods (separate or simultaneous optimization), we find an unstable behaviour of the calibrated mean reversion parameters. We are now fitting to 2 co-terminals, that is, 20 swaptions, however not all the swaptions are independent d.o.f. Considering that we are using 10 + 4 model parameters, though constrained, it appears that we do not have enough information to fix their values in a stable way suitable for practical use. Furthermore even if we allow the calibration to run in this manner, we do not observe a drastic improvement due to the introduction of the time-dependent mean reversion. We conclude that the new d.o.f. coming from the mean reversion are superfluous when only 2 co-terminals are fitted and that it is safer to use a constant mean reversion.

#### 4.4.4 Conclusion

A significant improvement in the fitting quality can be observed when including the mean reversion in the calibration procedure. Tests show that a constant mean reversion together with a time-dependent volatility can be obtained in a stable way while preserving a good quality of fit. It appears difficult and not particularly profitable to optimize on a time-dependent mean reversion. Simultaneous calibration of the constant mean reversion together with the volatility also brings a clear improvement of the fit while retaining stability.

While the 2 chosen co-terminals are approached with a reasonable accuracy by the model implied volatilities, we do not have any control on the rest of the swaption matrix. By focusing on 2 instruments per maturity, we may miss the other ones by quite a large amount, especially for early maturities, in a similar fashion to fig. 7. The hump in particular, is not captured at all.

### 4.5 All Instruments

Finally we consider the fit to all instruments in the swaption matrix.

It is often stated that Hull-White model, even with time-dependent volatility, is not able to fit the whole swaption matrix. What happens concretely is that, by trying to fit to all instruments, we lose control on which instruments will be fit correctly and which will not, see a typical example

of this in fig. 26 in the appendix. It may occur that only one instrument is fitted correctly, and the user would not be able to know in advance which one, which is quite a risky bet.

However, depending on the trade details and the purpose of the practitioner, it may be desirable to obtain a reasonable fit on the overall matrix. We have observed in the previous section that focusing on (a) particular co-terminal(s), even by optimizing on the mean reversion, has the disadvantage that other instruments may be seriously mis-priced. On the other hand, trying to optimize on a time-dependent mean reversion resulted in an unstable behaviour, which we attributed to the lack of information to constrain the mean reversion parameters.

It is therefore natural to wonder whether the information contained in the whole swaption matrix is suitable to stabilize a time-dependent mean reversion, and if so, how good the resulting fit would be. In this section we attempt to answer these questions. As method 2 did not seem to yield a significant runtime or stability improvement in the case of the calibration to 10Y and 20Y co-terminals, we will consider here only method 3, i.e. simultaneous calibration of the mean reversion and volatility in one multi-dimensional optimization.

First of all, in fig. 12 we compare the calibration targets over 2 years for the 3 strategies constant  $a$  and  $\sigma$ , constant  $a$  and time-dependent  $\sigma(t)$ , and finally both time-dependent  $a(t)$  and  $\sigma(t)$ .

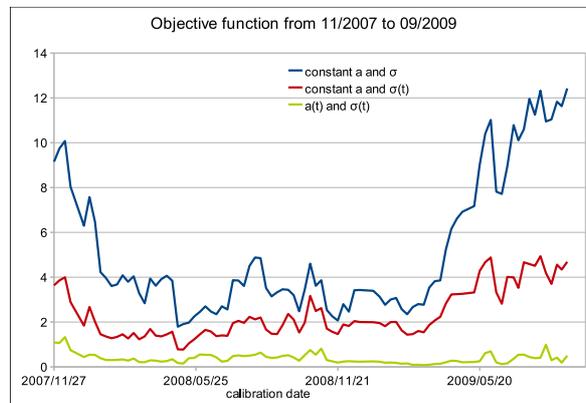


Figure 12: All instruments are fitted.

We can see that the time-dependence of the mean reversion gives us a spectacular improvement in the fitting quality.

It also appears that by calibrating to the whole swaption matrix we are less sensitive to Lehman crisis as can be seen from the shapes of the curves in the figure above, where we can no longer see the target spikes observed when calibrating to the co-terminals only. Whether this is a good feature or not is a matter of opinion. Being able to fit in a satisfactory way even in times of crisis may seem desirable as it avoids the danger of mis-pricing (or mis-hedging). On the other hand, a trader may hope that the model will be able to signal unusual market conditions by producing unexpected numbers, such as a bad fit quality.

As to the stability of the parameters over the test period, we find the evolutions in fig. 13 and 14 for the short-term and mid/long-term mean reversion and volatilities.

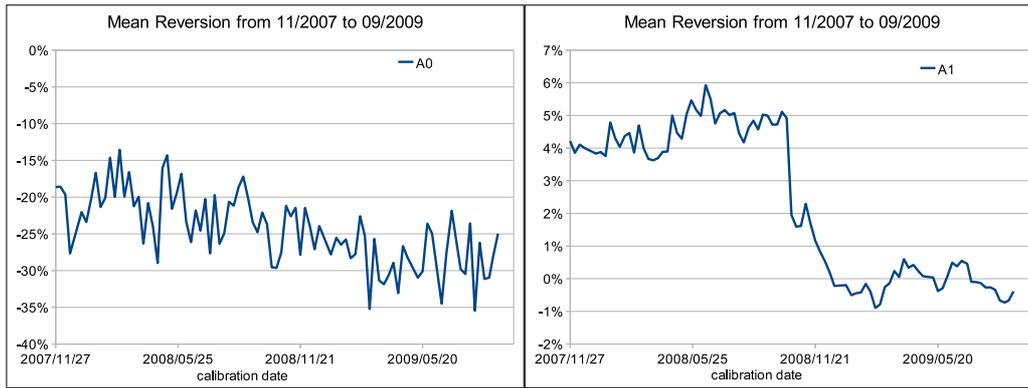


Figure 13: Parameters  $A_0$  and  $A_1$  for constraint  $\alpha = 0.1$

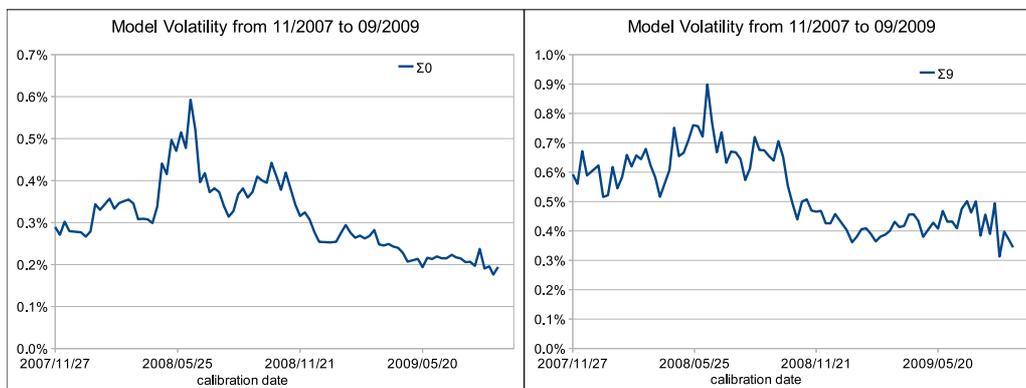


Figure 14: Parameters  $\Sigma_0$  and  $\Sigma_9$  for constraint  $\alpha = 0.1$

There does not seem to be any obvious instabilities. The behaviour of the parameters is similar to that observed for the constant calibrations, the mid/long-term mean reversion ( $A_1$ ) also showing the regime change after Lehman crisis.

To see in more details how the swaption implied volatility curves are fit, we will take two dates, roughly one year before Lehman crisis on the 2007/11/27 (fig. 15) and one year after the shock on 2009/09/08 (fig. 16)

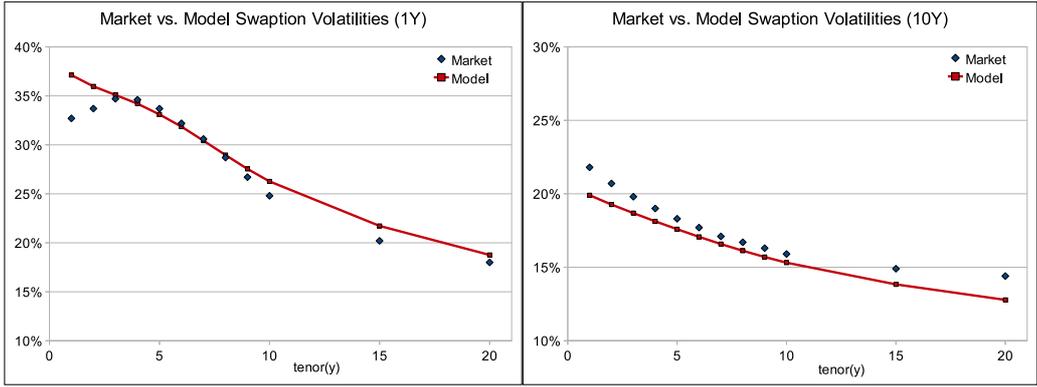


Figure 15: Market and Model Swaptions on the 2007/11/27 for  $\alpha = 0.1$

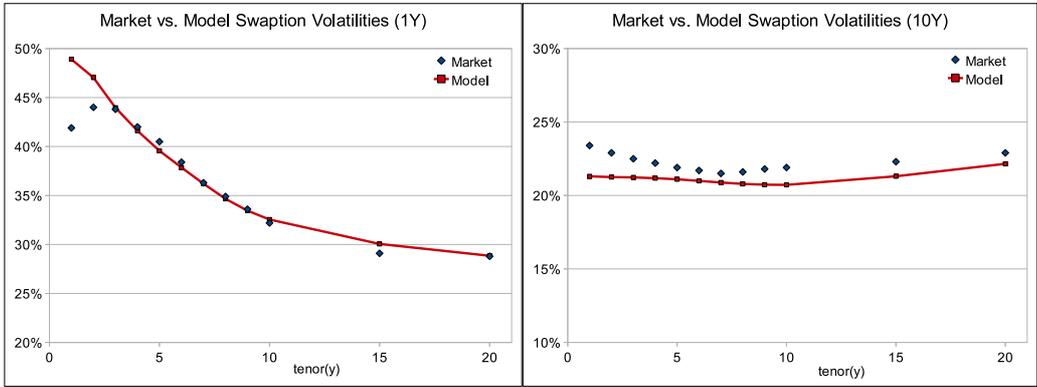


Figure 16: Market and Model Swaptions on the 2009/09/08 for  $\alpha = 0.1$

The quality of fit is greatly improved compared to the constant mean reversion case, and we no longer wildly mis-price any of the instruments, while retaining the stability of the parameters.

Finally, for the reference, we provide a similar result with a relaxed constraint for the time-dependent volatility: we take  $\alpha = 0.5$  now, and show the result for the calibration on the 2007/11/27 in fig. 17 below, while the result of the 2009/09/08 is left to the appendix in fig. 28.

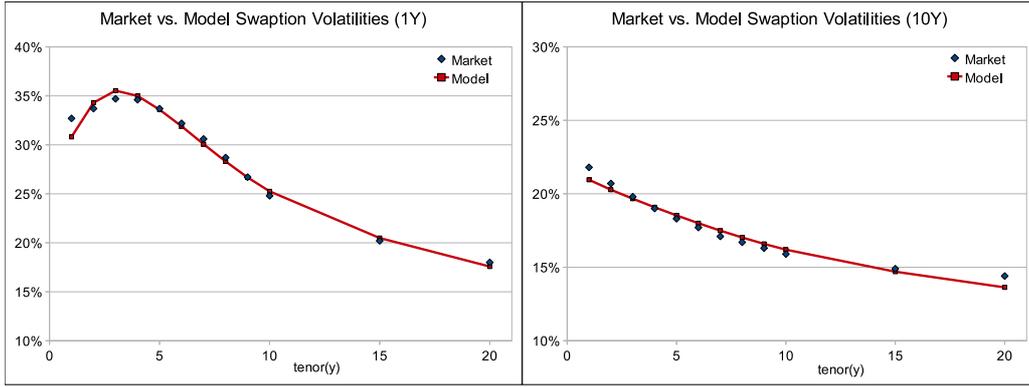


Figure 17: Market and Model Swaptions on the 2007/11/27 for  $\alpha = 0.5$

We can see that thanks to the new freedom we could fit the swaption even more accurately. We obtain now an excellent fit to the whole matrix, including the hump, while the parameters have the following evolution over the two year testing period

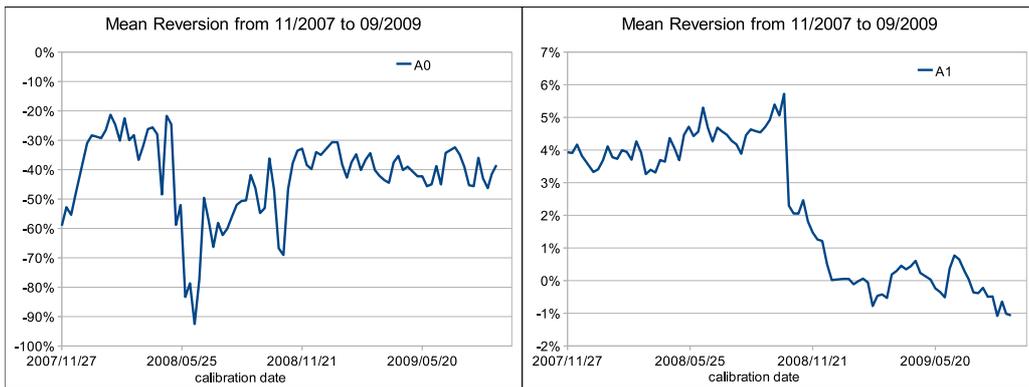


Figure 18: Parameters  $A_0$  and  $A_1$  for constraint  $\alpha = 0.5$

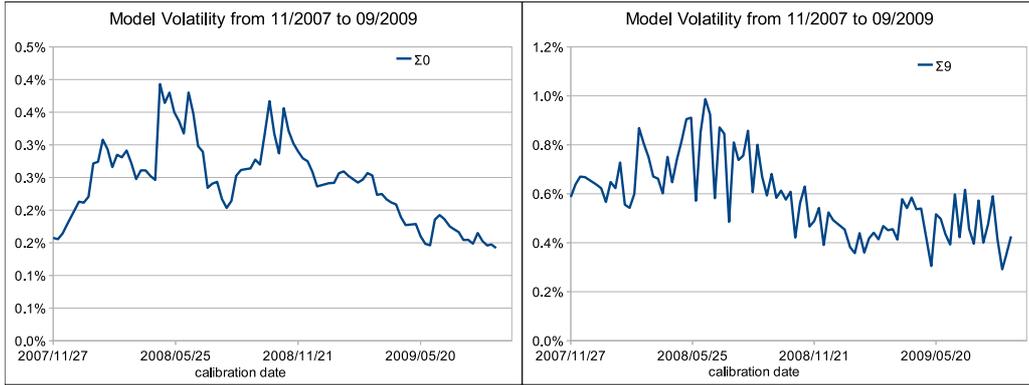


Figure 19: Parameters  $\Sigma_0$  and  $\Sigma_9$  for constraint  $\alpha = 0.5$

Unsurprisingly, they seem to bear wider moves than for the stronger constraint  $\alpha = 0.1$ , although how much stability has been lost remains ambiguous, as it is not clear how to measure it objectively. We do not see here evidence of over-fitting as can be observed in other calibration scenarios such as the unconstrained cubic spline or the calibration of a mean reversion to only one co-terminal, see fig. 24 and fig. 25.

#### 4.5.1 Conclusion

There seems to be enough information in the swaption matrix to obtain a stable estimation of a time-dependent mean reversion together with a time-dependent volatility. The resulting fit can be excellent over the whole swaption matrix without evidence of over-fitting.

### 4.6 Model and Market Evolution: analysis at Lehman Crisis

Although a perfect mirroring of the market evolution in terms of model parameters cannot be hoped for, due to limitations of Hull-White model as mentioned in the introduction, we have observed during this 2 year testing period a clear change in the mean reversion, visible across all types of calibration and parameter control strategies. As this behaviour is unambiguously exhibited by the model and its calibration methods, we can be confident that it is more than just a numerical coincidence and bears instead a deep meaning.

Recall that at the time of Lehman crisis, at the fall of 2008, the mean reversion suddenly drops to lower values, and does not return to its pre-shock values (at least until the end of the testing period in fall 2009). This can be observed for example in fig. 8 for constant mean reversions, or on the mid/long-term parameter  $A_1$  of the time-dependent mean reversion in fig. 13. At the same time, the model volatility does not seem to react in a significant way, such that one can say that Lehman crisis has been essentially absorbed in a mean reversion regime change.

As this change is obvious on the model side, we would like to know what is its market counterpart. In other words, we are trying to figure out what is or what are the features of the swaption matrix and the yield curve that have drastically changed at and after Lehman shock, and why they affected the model parameters in such a way.

It is a rather difficult question, as there are multiple degrees of freedom involved on both parts (market and model) and as the modelled variable, the short-rate, is not observable directly. However

we believe it is possible to draw a few conclusions.

First of all, we would like to separate, if possible, the influence of the swaption matrix from that of the yield curve. To this end, we go back to the 10Y and 20Y co-terminals calibration, with constant mean reversion and volatility for simplicity. We again calibrate over the 2 year testing period, except that this time, we fix the swaption matrix to its value on the first day, and only change the yield curve. We obtain the calibration result in fig. 20, with the legend "VC – FS". Then we can fix the yield curve on the first day, and change only the swaption matrix. This result is displayed in fig. 20 under the legend "FC – VS". And finally, we can perform a "real" calibration where both the curve and the swaptions vary. This has already been done and is represented in fig. 20 by the legend "VC – VS".

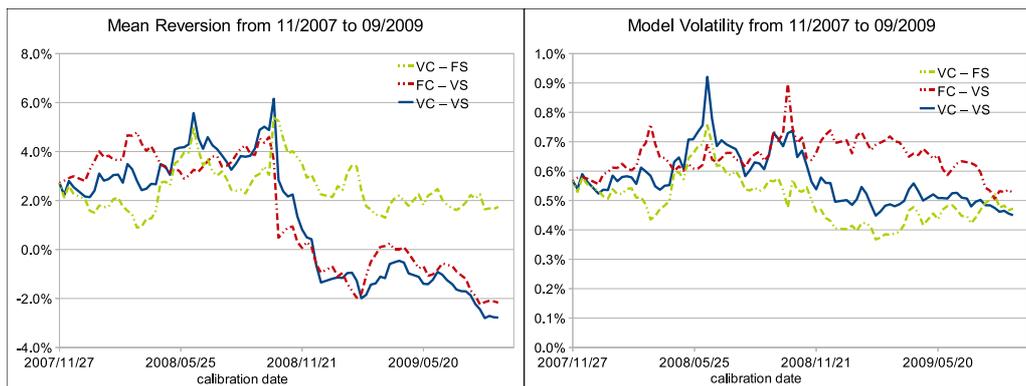


Figure 20: Difference in influence of the yieldcurve and the swaptions on the model parameters.

To sum up,

- VC – FS: the curve varies but the swaptions are fixed. Although there is a reaction of the parameters when Lehman crisis is triggered, this effect is not propagated afterwards and the parameters go back to similar values as before the shock.
- FC – VS: the curve is fixed but the swaptions varies. In this case we observe a behavior very similar to that of the real calibration, VC – VS, in that the mean reversion falls at the crisis and stays low afterwards.
- VC – VS: both the curve and the swaptions vary. This is the real calibration.

From all this we conclude that the regime change in the mean reversion has been caused mostly by the evolution of the swaption matrix<sup>2</sup>.

We display then the evolution of 4 different swaption implied volatilities (market) in fig. 21.

<sup>2</sup>Assuming that the "swaption  $\times$  yield curve" cross influence is negligible.

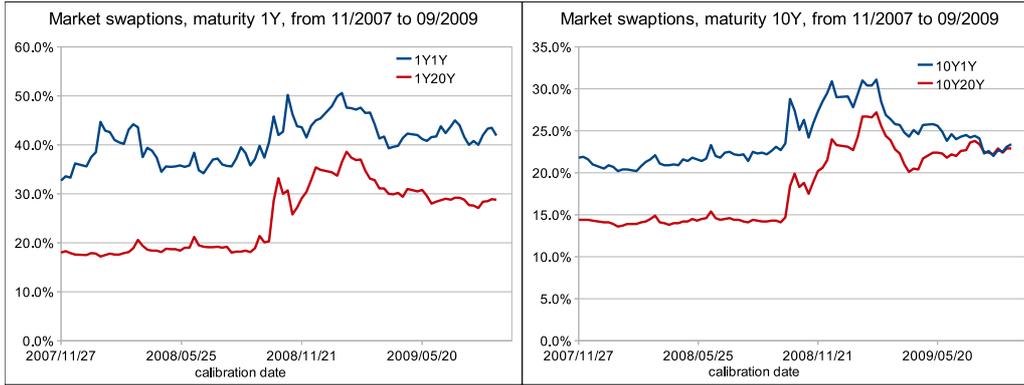


Figure 21: Evolution of JPY swaption implied volatilities

We observe two basic effects here. First of all, a level change appearing at the shock time and continuing afterwards. Second, a change in the relative values of the implied volatilities for the same maturity. This change is larger and most obvious for the 10Y swaptions, but is also visible at the 1Y maturity.

As we have discussed in section 4.1, fig. 5, a strong change in the ratios of implied volatilities at fixed maturity cannot be accounted for by a change in the model volatility, but rather in the mean reversion. Furthermore, since we observe in fig. 21 that the market implied volatilities have become flatter in the tenor direction, we deduce that this effect should be accounted for by a reduction of the value of the mean reversion, going possibly to negative values. This is indeed what is observed and what we have called a "regime change". It constitutes one more argument in favor of the consistency of the calibration methods proposed in this work.

#### 4.7 Bermudan Swaptions

Finally we address an issue that we have not dealt with until now, i.e. the fixing of the mean reversion in the case of a local calibration to one co-terminal. As we have shown in section 4.3, it is not safe to optimize on it in this case. This means that the mean reversion has to be fixed by the user, leading to an ambiguity on its value. Moreover, the subsequently calibrated volatility will depend on the chosen mean reversion value, so one may wonder what can be the influence of this a priori ad hoc choice on products other than simple vanilla options.

Here we fix 3 different (constant) mean reversions and calibrate the time-dependent model volatility to the 10Y co-terminal over the 2 year testing period. As the volatility is time-dependent, we obtain very close fits to the European swaptions whatever the mean reversion is. For the 10Y Bermudan swaptions, we find the results in fig. 22.

As was also documented in [4], we observe a strong influence of the mean reversion on the Bermudan swaption prices, even though the relevant European swaptions are almost perfectly fit. This means that a trader can use this extra freedom to impose a view on the Bermudan swaptions. Alternatively, this can be viewed as an extension of the calibration process, in which Bermudan swaption information is also included.

This freedom may be considered as desirable and may motivate the choice of the local calibration method instead of the global one. Indeed, we have seen in this case that calibrating the mean

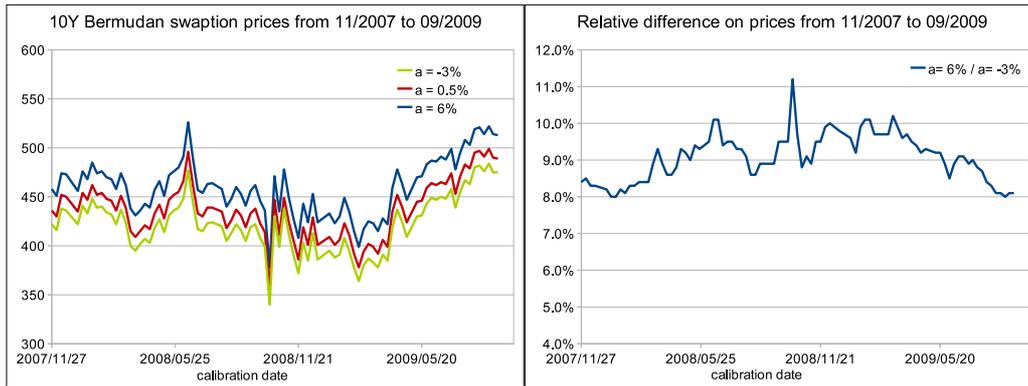


Figure 22: 10Y Bermudan Swaption Prices (non-call 3Y) for mean reversions 6%, 0.5% and  $-3\%$ .

reversion is preferable and therefore there is no more freedom in the model to impose a view on other products.

Note that this is true for the particular parametric form that we have chosen for the mean reversion, with its corresponding number of degrees of freedom (4). One may think that allowing more degrees of freedom in the mean reversion and optimizing only on some of them using the European swaption information may lead to both a stable calibration procedure with good fit to the European swaptions while still allowing to impose a view on the Bermudan swaptions.

## 5 Conclusion

The achievements of this work are two-fold. On the one hand, we provided a detailed documentation and numerical examples about the calibration methods of Hull-White model with time-dependent parameters, as we believe this was missing from the literature. We hope that the readers will find it useful as a starting point to implement their own calibration and that it will spare them the trouble of tedious trials and errors.

On the other hand, we showed that some common negative opinions and fears about the model are not necessarily justified. It is widely stated, for instance, that Hull-White model cannot fit the swaption matrix well, and that introducing time-dependent parameters to improve the fitting quality results in unstable behaviours. We believe this work has proved this is not the case, provided a suitable calibration strategy is adopted. In particular, we showed that the fully time-dependent model can achieve excellent fit in a stable manner. Furthermore, we provided some means, such as the "SMM" approximation of the implied volatilities, to interpret the behaviour of these parameters w.r.t. market moves. In particular, we showed that the model can provide interesting insights in the analysis of Lehman crisis in the fall of 2008.

It is not our purpose here to single out one best method. Instead we show that several calibration methods, such as local or global, with constant or time-dependent parameters, are acceptable, and that the choice depends on the user preferences as to fitting quality, runtime and ease of implementation. More specifically, it is clear that time-dependent parameters offer a better fitting quality, which comes at the cost of a slightly more complicated implementation of nested analytical integrals. We showed how to avoid the dangers of over-fitting, by using parametric forms with further constraints. However this method requires, in most cases, multi-dimensional constrained optimizations, representing an extra implementation and manipulation difficulty. In these tests we have used a customized version of the simulated annealing algorithm, which, while easily dealing with constraints, converges quickly enough for most practical purposes. In the particular case of a calibration with constant mean reversion and volatility though, the 2-dimensional optimization can be broken down to two 1-dimensional optimizations due to a property (approximate) of the model implied volatilities, captured by SMM approximation. For a user favoring ease of implementation and speed over fitting accuracy, a quasi-instantaneous method is thus available in which the mean reversion can be estimated. Note that this approximation may be made more accurate using the ideas of [1]. The generalization to time-dependent parameters, and its application to the calibration procedure, are currently under investigation in [10].

We believe that a number of open questions deserve further investigation in the future. First of all, the notion of instability of the parameters is quite ambiguous. In order to be able to run calibration in an automated way and discard the unstable methods, it would be useful to have an objective criterion as to what constitutes an unstable set of parameters.

Furthermore, the parametric forms proposed here are only suggestions and by no means the only possibilities. It may prove valuable to investigate the behaviour of other parametric forms, with different degrees of freedom. As we have briefly mentioned in the case of the local calibration, leaving the constant mean reversion undetermined by the calibration to European instruments may be desirable. The parametric forms and constrained multi-dimensional optimizations proposed here provide a suitable starting point to generalize this idea to more flexible calibration schemes in which the information of other products, such as Bermudan swaptions, may be incorporated.

Finally, the methods described in this document have been tested in details in the case of JPY currency (and briefly for USD and EUR) for the two year period Nov. 2007 to Dec. 2009. Although this includes the stress case of Lehman crisis, we would find it interesting to perform similar studies over different periods including other shocks, and for different currencies, in order to gain a deeper

understanding of the model implications and possibly discover other regime changes in the market.

## Acknowledgments

We would like to thank T. Hayashi for providing us with convenient and efficient tools to access market data, and O. Lenoble for his remarks which helped improve the manuscript.

## Appendix

### A Integrals

The basic integral that appears in most intermediate equations and final closed forms is based on integrating only the mean reversion. We denote it by  $E(t)$ , and it has the expression

$$E(t) = e^{\int_0^t a(u)du}. \quad (30)$$

From it we obtain the  $B(t, T)$  function involved in the affine structure of the discount bond

$$B(t, T) = E(t) \int_t^T \frac{du}{E(u)} \quad (31)$$

for which the following properties hold

$$\frac{\partial}{\partial t} B(t, T) = a(t)B(t, T) - 1 \quad (32)$$

$$\frac{\partial}{\partial T} B(t, T) = \frac{E(t)}{E(T)} \quad (33)$$

$$B(t, S) - B(t, T) = \frac{E(t)}{E(T)} B(T, S). \quad (34)$$

The short-rate is integrated explicitly as

$$r(t) = \frac{E(s)}{E(t)} r(s) + \alpha(t) - \frac{E(s)}{E(t)} \alpha(s) + \frac{1}{E(t)} \int_s^t E(u) \sigma(u) dW^{\mathbb{Q}}(u) \quad (35)$$

$$\alpha(t) = f(0, t) + \frac{1}{E(t)} \int_0^t E(u) \sigma^2(u) B(u, t) du. \quad (36)$$

where  $0 \leq s \leq t$  and  $f(0, t)$  is the initial instantaneous forward rate. The variance of the short-rate is given by

$$V_r(s, t) = \frac{1}{E^2(t)} \int_s^t E^2(u) \sigma^2(u) du. \quad (37)$$

$r(t)$  can be negative with non-zero probability

$$P(r(t) < 0) = \mathcal{N}\left(-\frac{\alpha(t)}{\sqrt{V_r(0, t)}}\right) \quad (38)$$

where  $\mathcal{N}(x)$  is the normal cumulative distribution function.

For the function  $\theta(t)$  involved in the SDE (1) and ensuring exact replication of the initial curve, we find

$$\theta(t) = \frac{\partial}{\partial t}f(0, t) + a(t)f(0, t) + \frac{1}{2}\left(\frac{\partial^2}{\partial t^2}V(0, t) + a(t)\frac{\partial}{\partial t}V(0, t)\right) \quad (39)$$

$$V(t, T) = \int_t^T \sigma^2(u, T)du \quad (40)$$

$$\sigma(u, T) = \sigma(u)B(u, T). \quad (41)$$

where  $\sigma(u, T)$  is the volatility at time  $u$  of the discount bond with maturity  $T$ .

In the discount bond affine structure

$$P(t, T) = \exp\left(A(t, T) - B(t, T)r(t)\right) \quad (42)$$

the  $A(t, T)$  function is

$$A(t, T) = \ln \frac{P(0, T)}{P(0, t)} + B(t, T)f(0, t) - \frac{1}{2}B(t, T)^2V_r(0, t). \quad (43)$$

The variance of the bond ratio

$$V_p(t, T_F, T_P) = \int_t^{T_F} \left(\sigma(u, T_P) - \sigma(u, T_F)\right)^2 du \quad (44)$$

factorizes as

$$V_p(t, T_F, T_P) = V_r(t, T_F)B(T_F, T_P)^2. \quad (45)$$

The correlation between two short-rates  $r_1 = r(t_1)$  and  $r_2 = r(t_2)$ , with  $t_1 < t_2$ , is

$$\rho(r_1, r_2) = \frac{E(t_1)}{E(t_2)} \sqrt{\frac{V_r(0, t_1)}{V_r(0, t_2)}}. \quad (46)$$

Finally, other useful formulae include the expression of the instantaneous forward rate at  $t$

$$f(t, T) = f(0, T) + \frac{E(t)}{E(T)} \left(r(t) - f(0, t) + \frac{V_p(t, T)}{B(t, T)}\right) \quad (47)$$

and the relation between the discount bound and the short-rate

$$\int_t^T r(u)du = -\ln P(t, T) + \frac{1}{2}V(t, T) + \int_t^T \sigma(u, T)dW^{\mathbb{Q}}(u). \quad (48)$$

## B Simulated Annealing

There are many different ways of running a simulated annealing algorithm, all centered on the same idea of drawing the next point randomly. Our version is based on the following steps

1. Choose an initial point  $x_0$  and set running variable  $x_r = x_0$
2. Evaluate the function and record its value under variable  $f_r = f(x_0)$
3. Draw a random point  $x$  under a normal distribution of mean  $x_r$  and standard deviation  $\sigma_j$
4. Check if  $x$  satisfies the constraints. If it does, go to step 5. If it does not, go back to step 3.
5. Evaluate  $f(x)$
6. i) If  $f(x) < f_r$  then the new point is closer to the minimum. It is thus recorded :  $f_r = f(x)$  and  $x_r = x$ .  
ii) If  $f(x) \geq f_r$  then the new point does not represent an improvement. We can forget about it.<sup>3</sup>
7. Go to step 3 until stop condition is met. Minimum is at  $x_r$  with function value  $f_r$ .

If we have some intuition of where the minimum is, we should take a small standard deviation  $\sigma_j$ . The algorithm will then focus on the region around our initial estimate and thus converge faster. If however we have little information about where this minimum could be (and its unicity, and the existence of a global minimum and local minima, etc...), we should take a large  $\sigma_j$ , such that the algorithm will explore a large portion of the parameter space.

For the stop condition, we advise to perform a number of convergence tests to observe how many function evaluations are necessary to reach a satisfying degree of convergence and runtime. In practice, for a 10-dimensional optimization, a few thousands of evaluations may be necessary, but runtime on a single 3.0GHz CPU typically does not exceed 10s.

Finally, we introduce a "cooling down" mechanism realized by a decrease of the standard deviation  $\sigma_j$  with each iteration  $j$ , according to

$$\sigma_j = \sigma_0 * \exp(-\gamma \times \frac{j}{N-1}) \quad (49)$$

where  $N$  is the total number of simulations,  $\sigma_0$  is the initial standard deviation and  $\gamma > 0$  describes the speed of cooling down.

Note that near the end of the algorithm, the standard deviation is  $\sigma_{N-1} \approx \sigma_0 * \exp(-\gamma)$ . If  $\gamma$  is chosen such that  $\sigma_{N-1} \approx 0$ , all remaining draws will be extremely close to the mean. In other words, the algorithm has been "frozen". This is analogous to the idea of cooling down for which a temperature parameter  $T$  (that may be chosen as  $T = 1 - \frac{j}{N-1}$  here) is decreasing little by little until reaching 0 and is used in Metropolis algorithm to allow or not uphill moves,  $T = 0$  signifying that no more moves are allowed and thus the algorithm has reached its end.

It is difficult to say what the values of the algorithm parameters  $x_0, N, \sigma_0$  and  $\gamma$  should be in general. A few tests are required to find a compromise between a satisfactory convergence speed and the ability to find the global minimum as well as the independence of the result on the initial point  $x_0$ .

---

<sup>3</sup>Metropolis algorithm can be used to allow "uphill" moves and increase the ability to escape from a local minimum.

## C Other Numerical Results

### C.1 SMM vs. Jamshidian Method

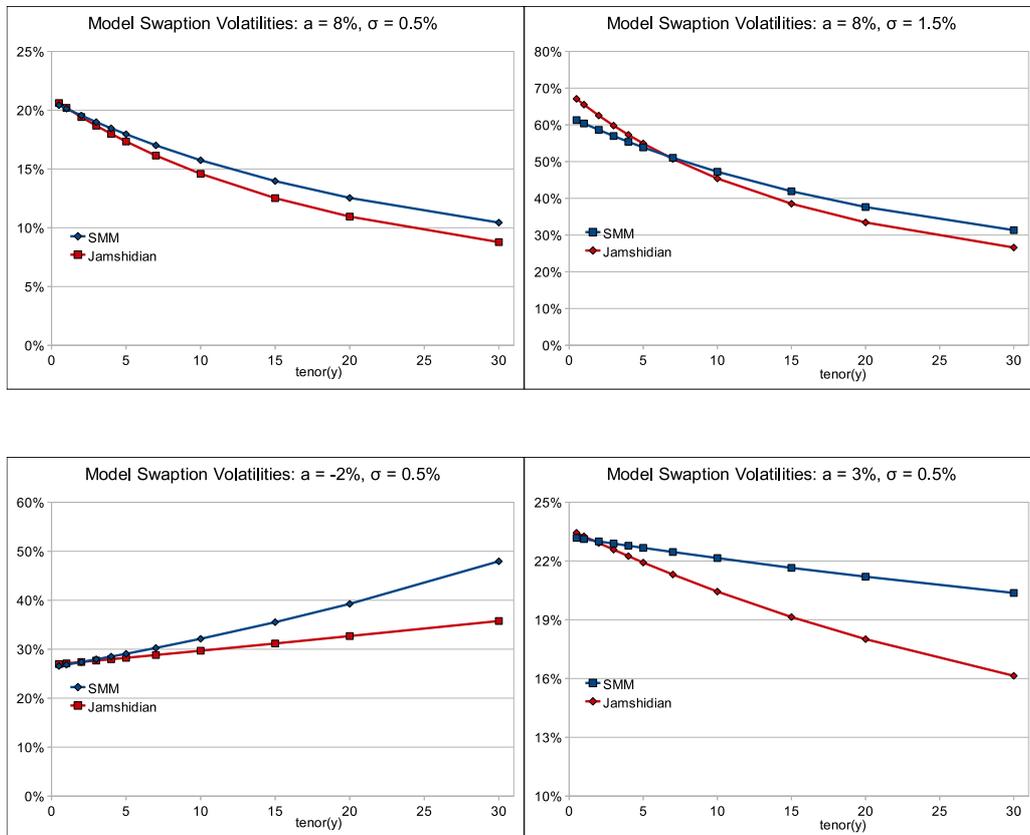


Figure 23: Swaption Implied Volatilities for various parameter values: Jamshidian Decomposition vs. SMM approximation.

## C.2 Other JPY calibration results

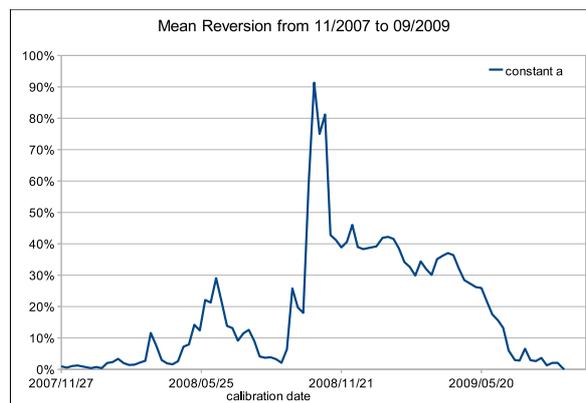


Figure 24: Instabilities in the mean reversion calibrated to 20Y co-terminal with Method 3.

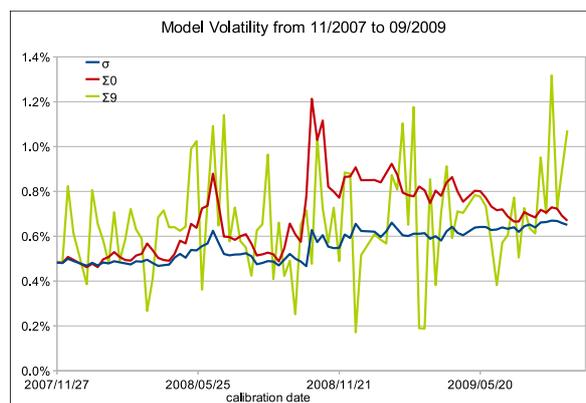


Figure 25: Model volatility calibrated to 20Y co-terminal at fixed mean reversion  
Displayed: constant  $\sigma$ , parameters  $\Sigma_0$  and  $\Sigma_9$  of  $\sigma(t)$ .

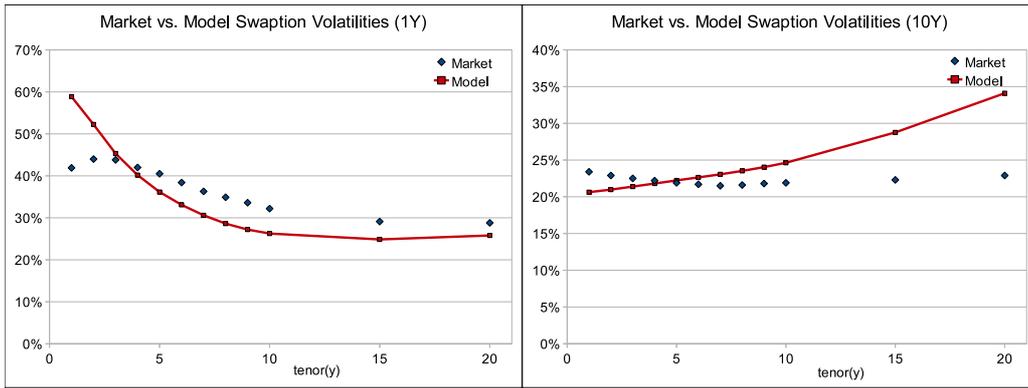


Figure 26: Market and Model Implied Volatilities on 2009/09/08.  
 All Instruments are calibrated, with constant  $a$  and time-dependent  $\sigma(t)$

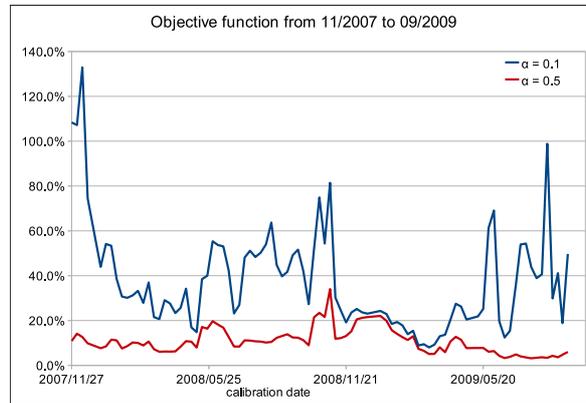


Figure 27: All instruments are calibrated. Mean reversion is time-dependent.

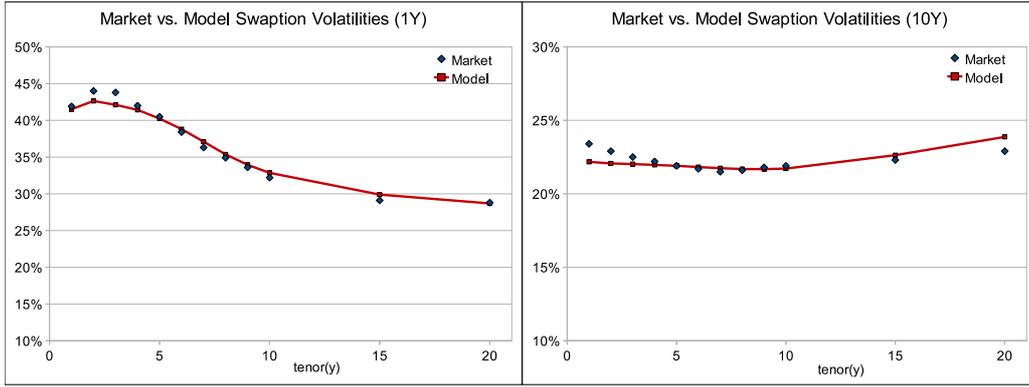


Figure 28: Market and Model Swaptions on the 2009/09/08 for  $\alpha = 0.5$ .

### C.3 Other currencies: USD

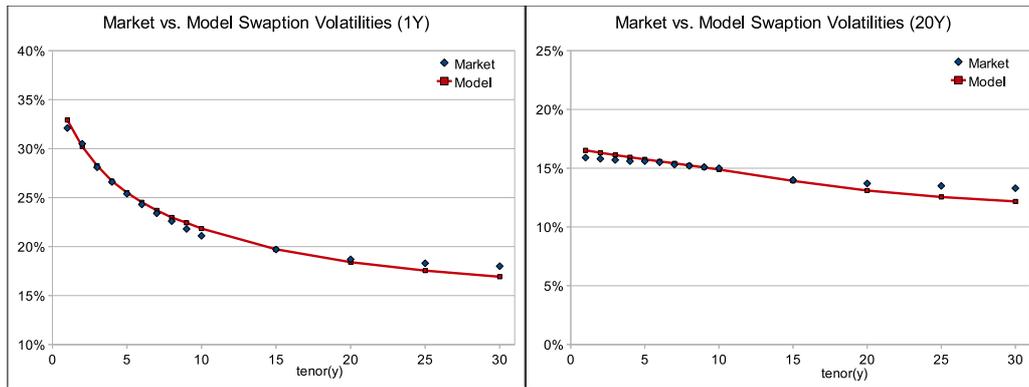


Figure 29: Market and Model USD Swaptions on the 2007/11/27 for  $\alpha = 0.1$   
 $A_0 = -1.2\%$ ,  $A_1 = 7.6\%$ ,  $\Sigma_0 = 1.27\%$ ,  $\Sigma_{13} = 1.42\%$

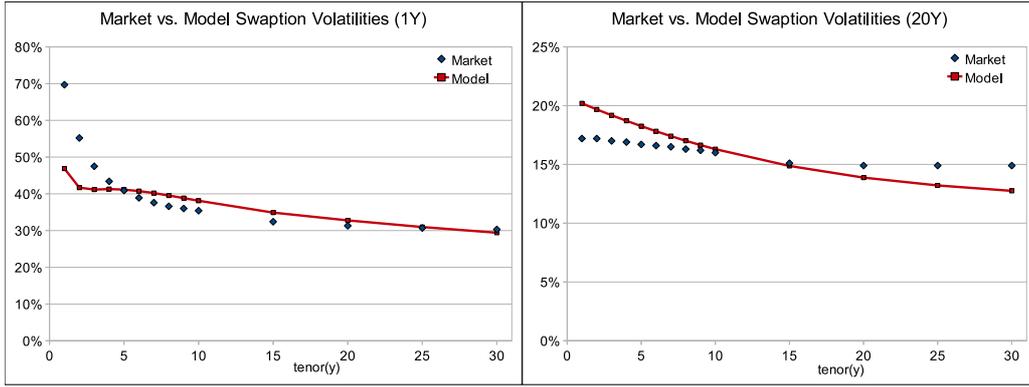


Figure 30: Market and Model USD Swaptions on the 2009/09/08 for  $\alpha = 0.1$   
 $A_0 = -41\%$ ,  $A_1 = 5\%$ ,  $\Sigma_0 = 0.52\%$ ,  $\Sigma_{13} = 1.44\%$

#### C.4 Other currencies: EUR

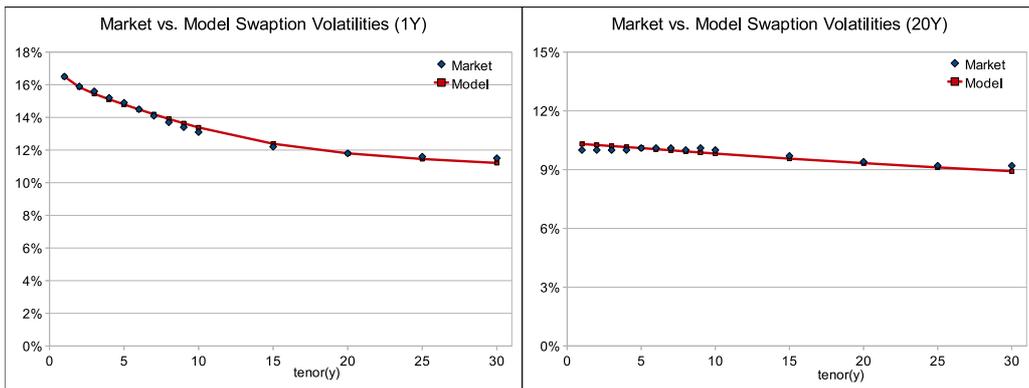


Figure 31: Market and Model EUR Swaptions on the 2007/11/27 for  $\alpha = 0.1$   
 $A_0 = -2.8\%$ ,  $A_1 = 6.1\%$ ,  $\Sigma_0 = 0.73\%$ ,  $\Sigma_{16} = 0.65\%$

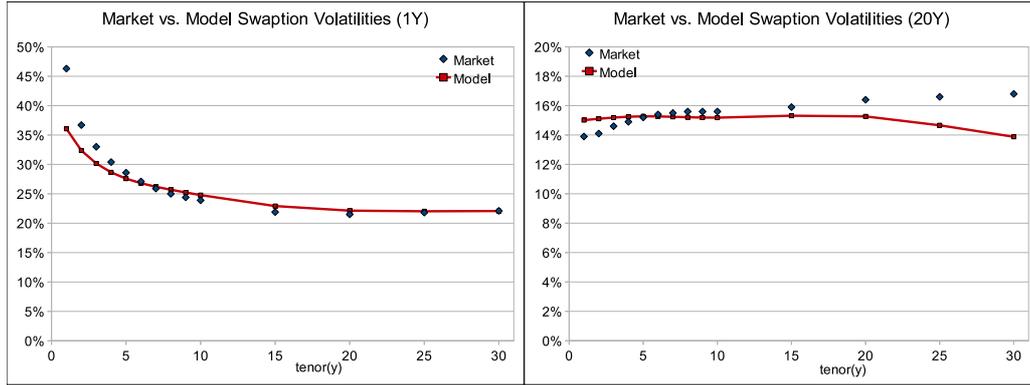


Figure 32: Market and Model EUR Swaptions on the 2009/09/08 for  $\alpha = 0.1$   
 $A_0 = -16.2\%$ ,  $A_1 = 3.2\%$ ,  $\Sigma_0 = 0.70\%$ ,  $\Sigma_{16} = 0.81\%$ .

## References

- [1] D. Schrage A. Pelsser. Pricing Swaptions and Coupon Bond Options in Affine Term Structure Models. *Mathematical Finance*, 4:673:694, 2006.
- [2] A. Brace, D. Gatarek, and M. Musiela. The Market Model of Interest Rate Dynamics. *Mathematical Finance*, 7:127–156, 1997.
- [3] D. Brigo and F. Mercurio. Interest rate models - theory and practice (with smile, inflation and credit). 2006.
- [4] P. J. Hunt and J. E. Kennedy. *Financial Derivatives In Theory And Practice*. John Wiley & Sons, revised edition, 2005.
- [5] A. White J. Hull. Pricing interest-rate derivative securities. *The Review of Financial Studies*, 3(4):573:592, 1990.
- [6] A. White J. Hull. Using Hull–White Interest Rate Trees. *The Journal of Derivatives*, 3(3):26:36, 1996. Spring.
- [7] C. White J. Hull. The General Hull-White Model and Super Calibration. *Financial Analysts Journal*, pages 34–44, November/December 2001.
- [8] F. Jamshidian. An Exact Bond Option Pricing Formula. *The Journal of Finance*, 44:205–209, 1989.
- [9] F. Jamshidian. Libor and Swap Market Model and Measures. *Finance and Stochastics*, 1:293–330, 1997.
- [10] T. Wong S. Gurrieri. A New Separability Property of the Hull–White Short Rate Model. *Working Paper*, 2010.