FX Column: How to Start a Stochastic Volatility Process

Impact on Mixing Local and Stochastic Volatility

Andreas Weber and Uwe Wystup
MathFinance AG

12-Jul-2026

Table of contents

1 Introduction

The standard approach to using a stochastic volatility (SV) model for valuation of a derivative is to assume one deterministic level of volatility at time zero and define how volatility evolves over time as a random process. If this process is continuous, after a short time the stochastic volatility will still be quite concentrated around the initial level. Consequently, the generated implied volatility for very short expiries is still quite flat. As the example of mixture log-normal models show, at least two different levels of volatility are needed to generate convexity of the implied volatility smile (D. Brigo and Rapisarda 2002).

To fit also more pronounced short tenor convexity, either a random start of the volatility process or high levels of mean reversion speed and volatility of volatility (generating a more jump-like process) have been proposed. (See Jacquier and Shi (2016) and further references therein for the first approach and Mechkov (2015a) for the second.)

We illustrate these approaches in the context of calibrating time-dependent SV parameters. We then discuss the fitting to the implied volatility surface using a stochastic-local volatility (SLV) processes. Finally, we look at the impact on the mixing factor when mixing stochastic and local volatility models, a popular approach in the FX asset class.

1.1 Heston model

We will use a variant of the Heston model (Heston 1993) as stochastic volatility model, where the center of mean reversion (long term variance) is always 1 and the overall level of volatility is a factor in the spot equation. The model equations are

$$\begin{aligned}
\frac{dS_t}{S_t} &= \mu_t dt + \sigma_t \sqrt{v_t} dW_t^1, && (1.1)\\
dv_t &= \kappa_t(1 - v_t) + \xi_t \sqrt{v_t} dW_t^2, && (1.2)\\
dW_t^1 \cdot dW_t^2 &= \rho_t dt. && (1.3)
\end{aligned}$$

Factoring out the level of volatility and modeling only convexity (ξ) and skew (ρ) by the SV model is a technique used sometimes; for a reference and details see van der Zwaard (2016).

We will consider the parameters κ_t, ξ_t, ρ_t, and σ_t as piecewise constant, typically constant between tenors of the market. Only ξ_t, ρ_t, and σ_t will be calibrated, κ_t will be set to either a constant value of 2, or to some time dependent value, but in any case the value is fixed before calibration (both κ and ξ have an impact on the convexity). It could be discussed whether to consider σ_t as piecewise constant (instead of continuous). However, in the context of stochastic local volatility (see section 3) another calibration step will overwrite this level of volatility anyway.

We illustrate the examples using some EUR-USD market data set from 2021. For each tenor, five volatilities are quoted for at the money and the 10% and 25% call and delta equivalent strikes.

1.2 Slicewise Approach

In a first approach ("Slicewise") we calibrate the parameters to fit the 1W volatilities. Then we fix these parameters for the period up to one week and calibrate the parameters for the next period (between the 1W and 2W tenor) such that the 2W volatilities are fit as good as possible. This is iterated for all the tenors.

In the graphs we plot the implied volatilities from the model (Heston, slicewise calibration) together with the five market data input volatilities and a parametrized curve fitting the market data points at the according tenor (SVI, Gatheral (2004)).

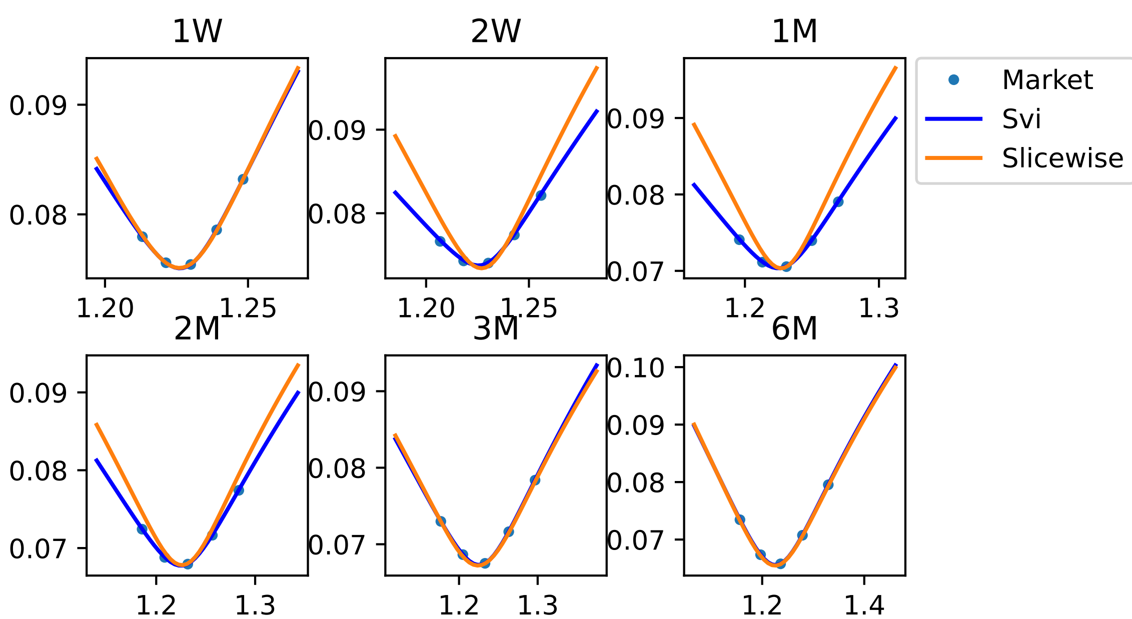


Figure 1.1: Slicewise calibration from short to long tenor

The first tenor is very well fitted. But for the following tenors the Heston smile shows a higher convexity than market smile.

Table 1.1: Slicewise calibration from short to long tenor

	Tenor	Kappa	Sigma	Xi	Rho
0	1W	2.0	0.0785	7.15	0.11
1	2W	2.0	0.0791	0.16	0.12
2	1M	2.0	0.0752	0.26	0.90
3	2M	2.0	0.0695	0.24	0.89
4	3M	2.0	0.0693	0.27	0.90
5	6M	2.0	0.0720	6.01	0.05
6	9M	2.0	0.0738	2.42	0.16
7	1Y	2.0	0.0745	6.48	0.08
8	2Y	2.0	0.0804	4.32	-0.03

Looking at the ξ of the calibrated parameters, we observe a high value for the first tenor, followed by values close to zero. It seems that the first ξ is very high and even a value of 0 for the following periods is not low enough to bring down the time average to fit the implied convexity for the time up to the following tenors.

Also for longer tenors we see a zigzagging pattern (alternations between high and low values).

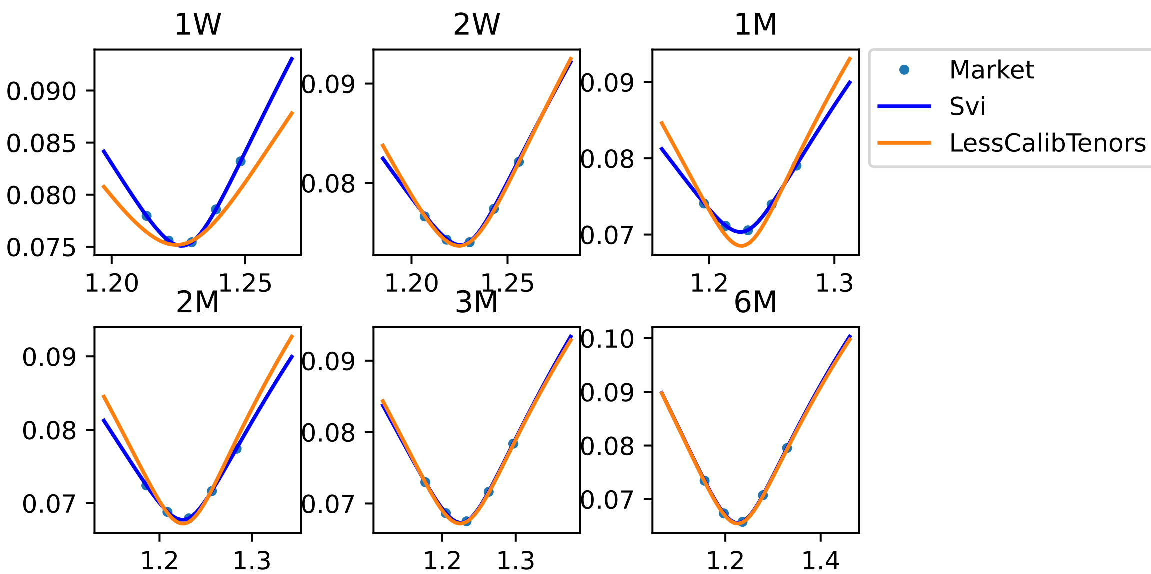
Aside from low quality fit for the tenors 2W to 2M, the parameters define a stochastic volatility model with almost no stochastic volatility during this period.

In the next section we will illustrate some approaches to avoid this pattern. Section 3 will add a local volatility term to the stochastic volatility model and illustrate that the additional term can compensate a lack of fit of the stochastic volatility model. Sections 4 and 5 will demonstrate that nevertheless for path dependent products the calibration of the stochastic volatility model will matter.

2 Some variations

2.1 Less calibration tenors

Let us select a subset of tenors for the calibration. Perhaps we can avoid the high value of ξ of the first tenor and leave more time until the next calibration tenor. For a subset 2W, 3M, 6M, 1Y, 2Y we obtain the following fit:



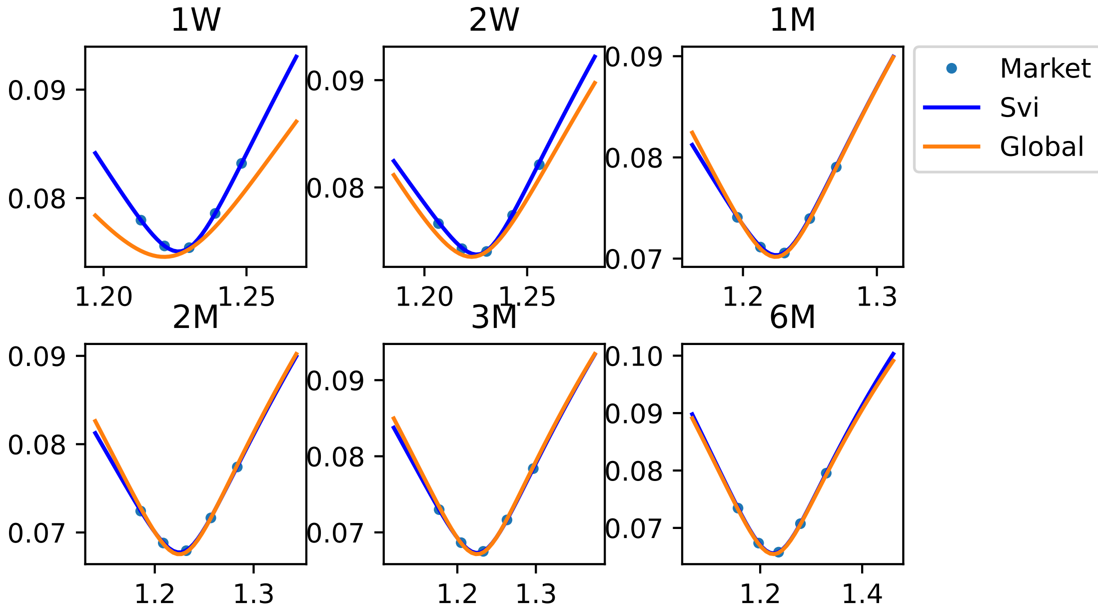
	Tenor	Kappa	Sigma	Xi	Rho
0	2W	2.0	0.0772	5.28	0.12
1	3M	2.0	0.0713	0.65	0.24
2	6M	2.0	0.0720	5.71	0.06
3	1Y	2.0	0.0740	4.00	0.11
4	2Y	2.0	0.0801	4.83	-0.03

We still observe a zigzagging pattern in the term-structure of ξ. The Heston smile for 1W is now too flat, the for 1M and 2M too steep. (The overall level for 1M could be corrected by calibrating σ for 1M (but not by calibrating ξ).

2.2 Global optimization with penalty for zigzagging

One approach to preventing zigzagging is to use a parametric function over time for at least ξ_t and to calibrate the parameters of this function.

Alternatively we use a here a global optimization for the Heston parameters with adding some penalty for high variation of the parameters in time. As we observe the zigzagging especially for the short tenors, we apply the global optimization for the first five tenors only and continue slicewise calibration thereafter.

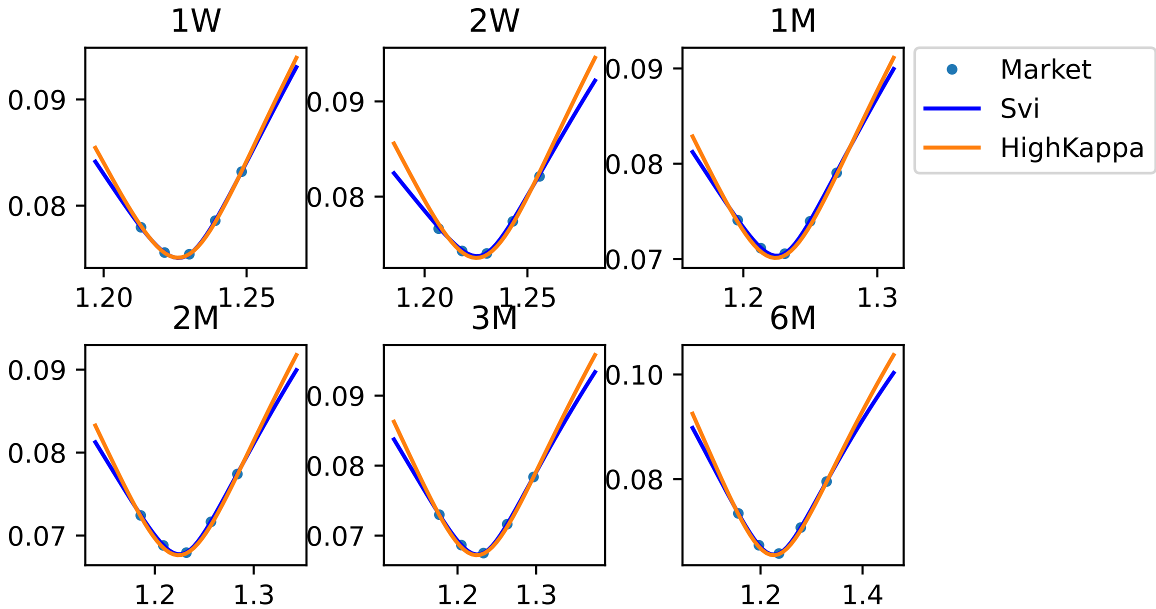


	Tenor	Kappa	Sigma	Xi	Rho
0	1W	2.0	0.0766	4.79	0.16
1	2W	2.0	0.0762	2.90	0.03
2	1M	2.0	0.0724	2.61	0.07
3	2M	2.0	0.0698	2.97	0.09
4	3M	2.0	0.0732	3.28	0.08
5	6M	2.0	0.0715	3.38	0.08
6	9M	2.0	0.0727	5.95	0.06
7	1Y	2.0	0.0749	3.66	0.16
8	2Y	2.0	0.0803	4.87	-0.04

The Heston smile for the fist two tenors is too flat (only slightly for 2W), but from 1M onwards the smiles are fitted well. More importantly, the parameters define now a real stochastic volatility model throughout all periods.

2.3 High kappa for short tenors

Another idea is to use a high kappa for short tenors as a high speed of mean reversion might allow to adjust the stochastic volatility in shorter time periods.

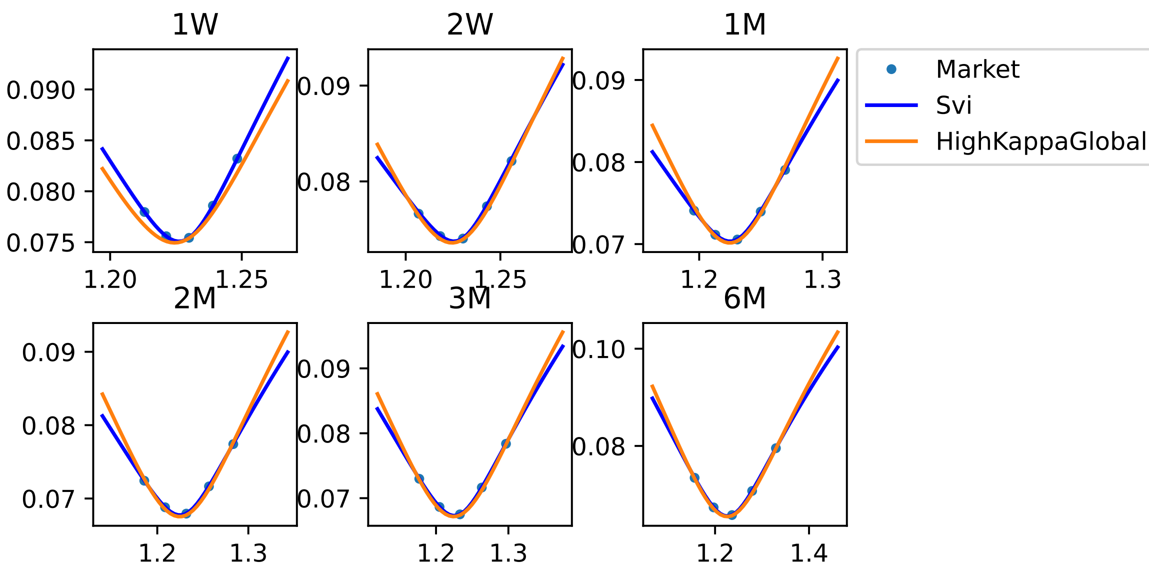


A graphical inspection shows that the implied smile of the SV now fits the market for short and long tenors rather well. But looking at the parameters we find again a zigzagging pattern and time periods with very low ξ.

	Tenor	Kappa	Sigma	Xi	Rho
0	1W	28.6	0.0786	8.81	0.11
1	2W	20.3	0.0764	0.61	0.80
2	1M	13.1	0.0714	6.75	0.06
3	2M	9.7	0.0698	6.76	0.09
4	3M	8.0	0.0734	8.43	0.07
5	6M	5.6	0.0726	7.70	0.08
6	9M	4.6	0.0733	10.26	0.07
7	1Y	4.0	0.0755	8.74	0.11
8	2Y	2.8	0.0804	6.41	-0.02

2.4 High kappa and global optimization

Here we combine both ideas.



	Tenor	Kappa	Sigma	Xi	Rho
0	1W	28.6	0.0776	7.43	0.13
1	2W	20.3	0.0765	6.67	0.08
2	1M	13.1	0.0725	6.57	0.06
3	2M	9.7	0.0696	6.80	0.09
4	3M	8.0	0.0727	6.87	0.09
5	6M	5.6	0.0726	8.35	0.07
6	9M	4.6	0.0735	9.97	0.07
7	1Y	4.0	0.0755	8.90	0.11
8	2Y	2.8	0.0803	6.39	-0.02

In this approach, the term-structure of ξ is moderately stable, the smile fits look acceptable, however, the high value of κ for the shorter tenors seem hard to justify: why should a stochastic process for the volatility have a significantly higher mean-reversion speed for the short term? To avoid this weak spot, we move on to the next section.

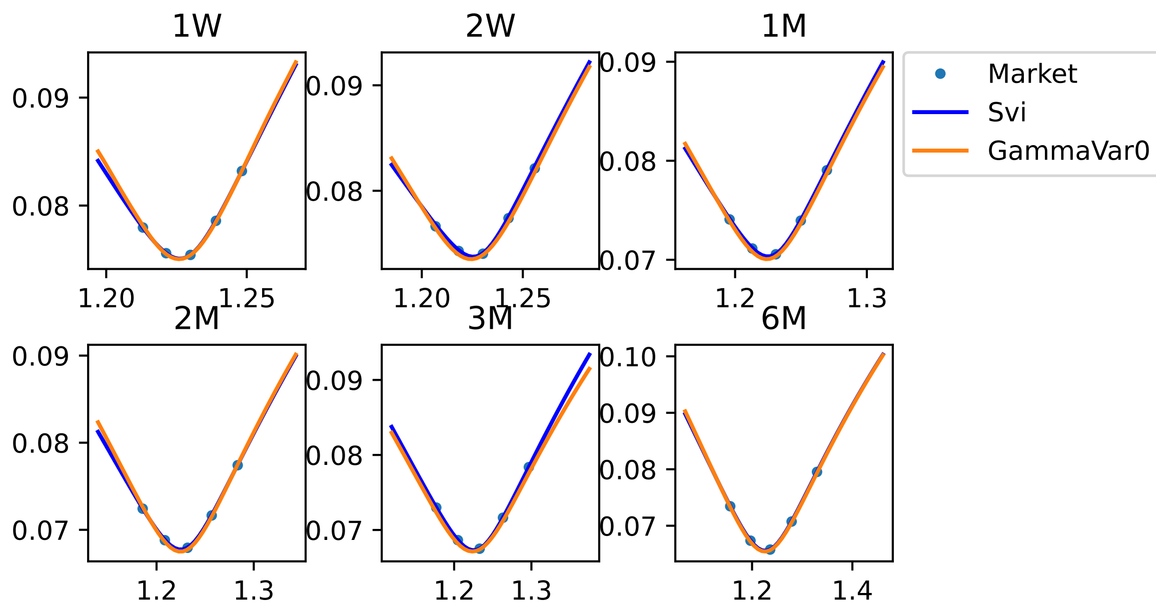
2.5 Starting the stochastic volatility process from a random level of volatility

With a constant (local) volatility in the spot equation of the Heston model we can produce only a flat smile (constant implied volatility). Convexity can be generated by the mixture of at least two different (stochastic or deterministic) levels of volatility. In a stochastic volatility model the volatility process typically starts from *one* deterministic point of volatility. After a short time the volatility process is still not diverged sufficiently to generate a real mix of different volatility levels.

As we observe a smile even for ON quotes, it seems likely, that there is uncertainty about volatility level also at time zero. This suggests that we should start the variance process from a random level.

2.5.1 Starting from a gamma distribution

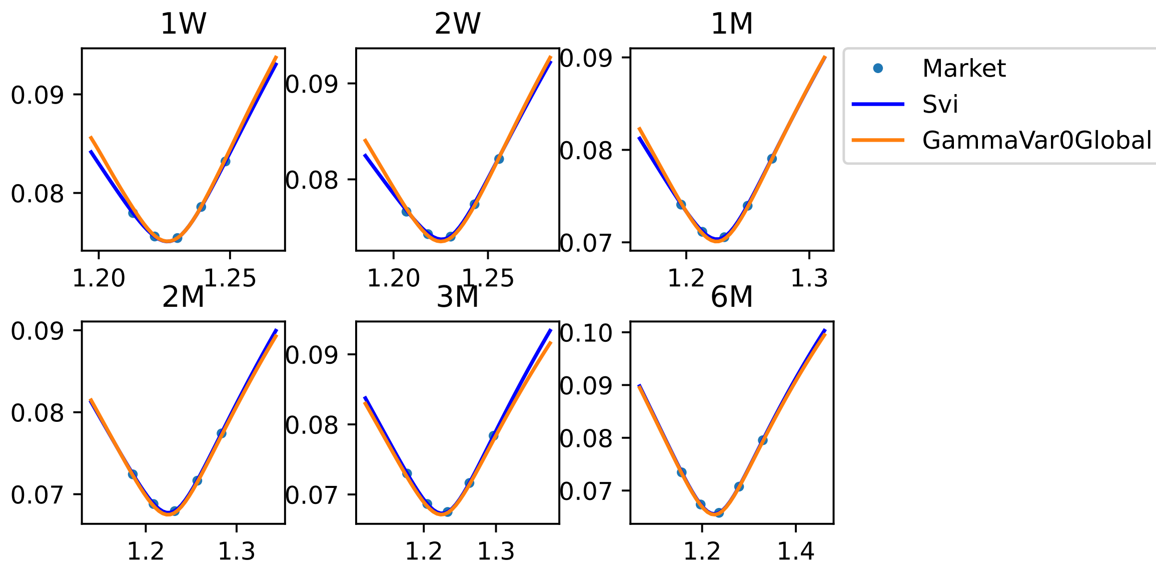
Mechkov (2015b) proposes to draw the starting variance for the Heston variance process from a gamma distribution, as the gamma distribution is the equilibrium distribution of the CIR process (the process describing the variance v).



	Tenor	Kappa	Sigma	Xi	Rho
0	1W	2.0	0.0790	1.54	0.52
1	2W	2.0	0.0758	1.26	-0.01

	Tenor	Kappa	Sigma	Xi	Rho
2	1M	2.0	0.0720	3.87	0.06
3	2M	2.0	0.0702	0.25	0.90
4	3M	2.0	0.0715	0.27	0.90
5	6M	2.0	0.0728	5.85	0.05
6	9M	2.0	0.0738	2.58	0.15
7	1Y	2.0	0.0745	6.35	0.09
8	2Y	2.0	0.0804	4.35	-0.03

While the smiles are fit well, and the level of ξ for the 1W tenor is much lower now, we still see a strong zigzagging pattern and a low levels for ξ during some time periods. Hence we will combine the method with the global optimization and obtain the following result.



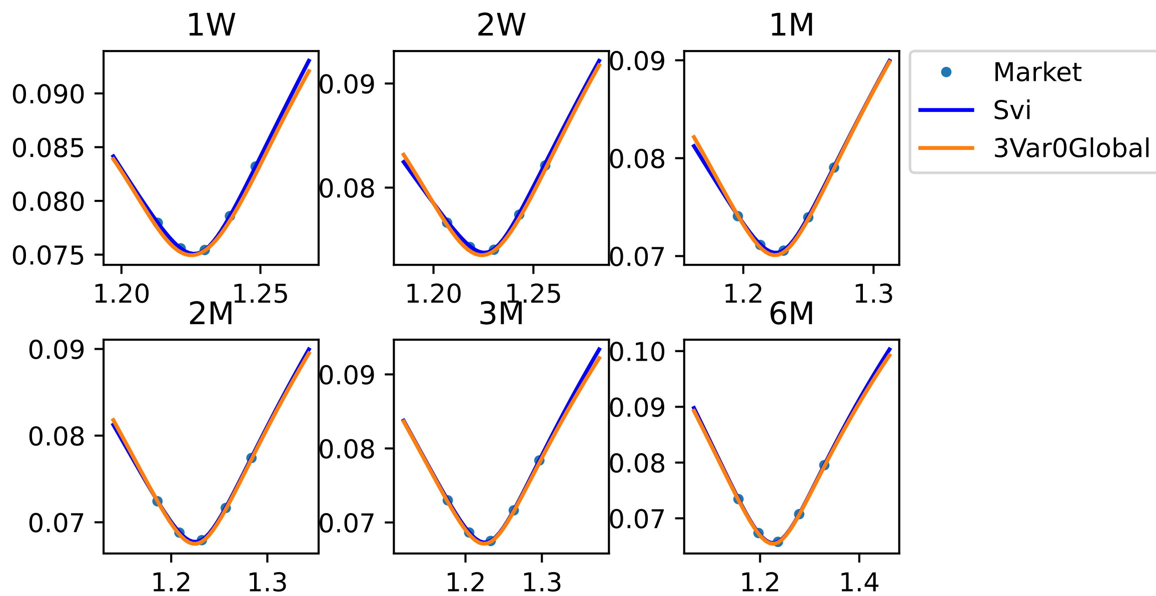
	Tenor	Kappa	Sigma	Xi	Rho
0	1W	2.0	0.0792	2.19	0.37
1	2W	2.0	0.0763	2.38	0.02
2	1M	2.0	0.0721	2.55	0.07
3	2M	2.0	0.0694	2.58	0.10
4	3M	2.0	0.0723	2.46	0.11
5	6M	2.0	0.0725	4.76	0.06
6	9M	2.0	0.0732	4.10	0.09
7	1Y	2.0	0.0745	5.40	0.11
8	2Y	2.0	0.0804	4.60	-0.03

2.5.2 Starting from a few discrete levels of volatility

Mixing two flat volatility models with different levels of volatilities already can generate an implied smile. By choosing different sets of levels, the convexity of the smile can be varied, see (MLV, Wystup (2020))

Rather than drawing the initial volatility from a continuous distribution, we will now select it from a discrete set of levels. This is done in such a way that, using a mixing model (i.e. mixing flat volatility models), we can generate the 1W convexity.

For the example data set, we end up choosing the initial variance for the Heston model from the set $\{0.51, 0.84, 1.65\}$ with equal probability.



	Tenor	Kappa	Sigma	Xi	Rho
0	1W	2.0	0.0781	2.92	0.27
1	2W	2.0	0.0758	2.63	0.02
2	1M	2.0	0.0719	2.61	0.08
3	2M	2.0	0.0693	2.75	0.08
4	3M	2.0	0.0724	2.74	0.11
5	6M	2.0	0.0720	4.30	0.06
6	9M	2.0	0.0729	4.73	0.08
7	1Y	2.0	0.0745	4.84	0.12
8	2Y	2.0	0.0803	4.70	-0.03

We conclude that in the combined approach the smiles are well fit, the Heston parameters reasonable in size and non-oscillating. This could be the end of the column, if we didn't have to deal with exotics. We move on.

3 Leverage function

Another approach to improve the fitting to an implied volatility surface is to extend the SV model with an additional factor in the equation describing the spot process via

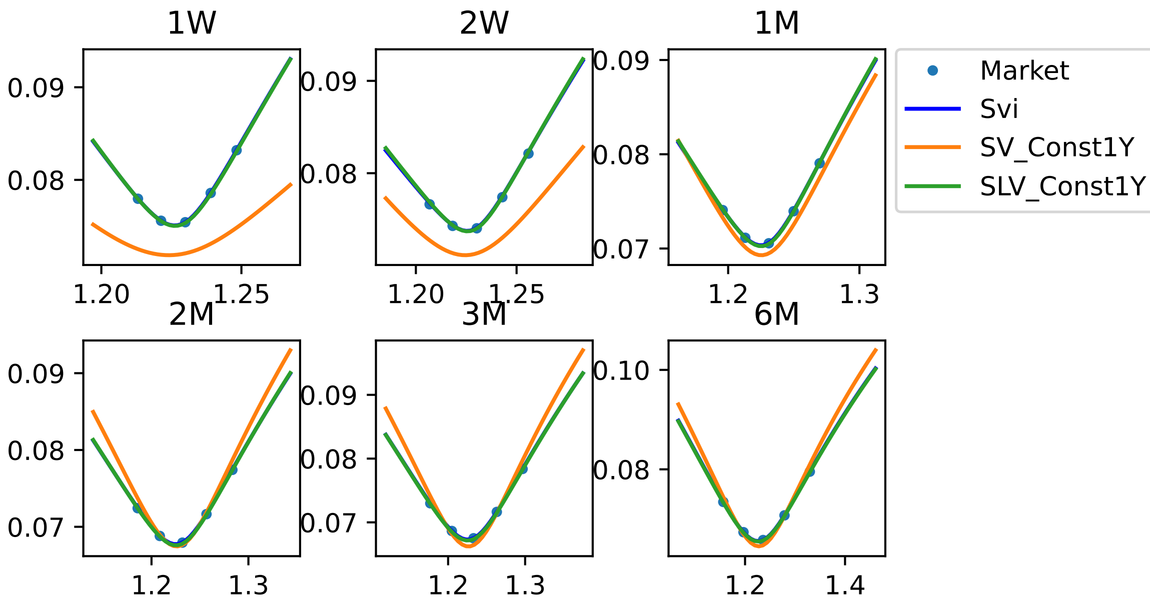
$$\frac{dS_t}{S_t} = \mu_t dt + \sigma_t l(S_t, t)\sqrt{v_t} dW_t^1 , \tag{3.1}$$

where $l(S_t, t)$ is called the *leverage function*. In case of a deterministic $v = 1$ and setting $\sigma_t l(S_t, t)$ equal to the Dupire local volatility, the implied volatility surface is fitted exactly, the model becomes a local volatility model. If v is stochastic, the model is called stochastic-local volatility model (SLV).

3.1 Constant SV parameters

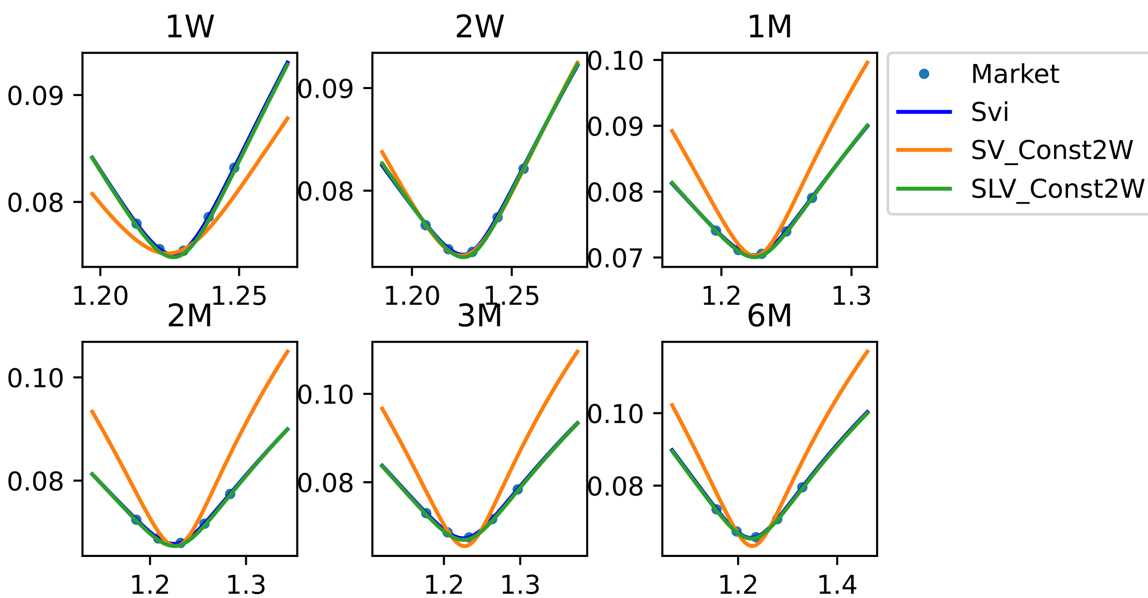
Even if we assume constant Heston parameters, the leverage function can be calibrated in such a way that the SLV implied smiles fit the market for all tenors, as we show in the following two examples.

We calibrate the Heston parameters to fit the 1Y smile only and use them as constant parameters for all time periods.



	Tenor	Kappa	Sigma	Xi	Rho
0	1Y	2.0	0.0729	3.78	0.09

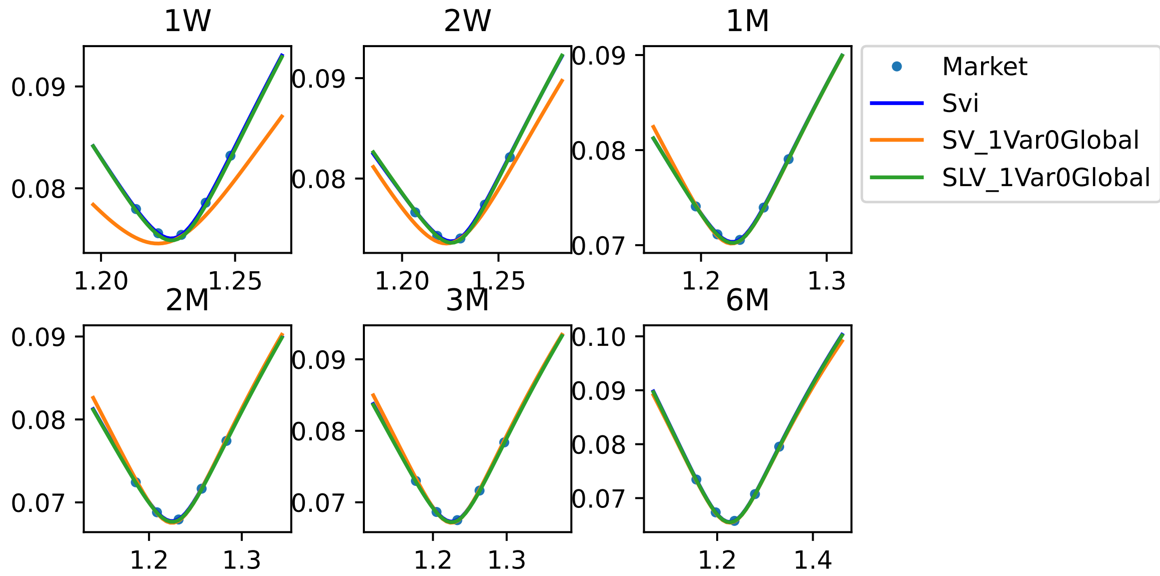
In the next example we assume constant Heston parameters, calibrated to the 2W smile.



	Tenor	Kappa	Sigma	Xi	Rho
0	2W	2.0	0.0772	5.28	0.12

3.2 Global with one starting level of variance

In this example we use the global calibration for the short tenors with one deterministic level of initial variance.



The approach with the leverage function allows to assume a wide range of stochastic volatility parameters while still fitting the implied volatility surface. However, to only fit the implied volatility surface the Dupire local volatility model is sufficient, without the need for stochastic volatility.

In the next section we will illustrate the impact of the model choice (calibration of the SV parameters) to a path dependent claim.

4 Model price of one touch contracts under different SV calibration methods

As an example for a path dependent derivative, we consider a one touch. The payoff is one USD, if at any time until expiry the EURUSD spot reaches a pre-defined barrier level.

In the *mustache* graphs below we plot the difference between a model price and the Black-Scholes price (assuming a flat constant volatility with the level of the at-the-money volatility) for a series of barrier levels. The differences are plotted against the Black-Scholes Theoretical Values (TV) (instead of the barrier levels); see (Mustache, Wystup (2019)).

The dotted lines show the prices under the specified SV model and calibration, the solid lines show the prices under the corresponding SLV model.

We exhibit the following models:

- 3Var0Global: Three discrete initial variances (Section 2.5.2)
- 1Var0Global: One initial variance (Section 2.2)
- HighKappaGlobal: Higher Kappas for short tenors (Section 2.4)
- Slicewise: Slicewise calibration (Section 1.2)
- Const1Y: Constant parameters, fitted to 1Y smile (Section 3.1)

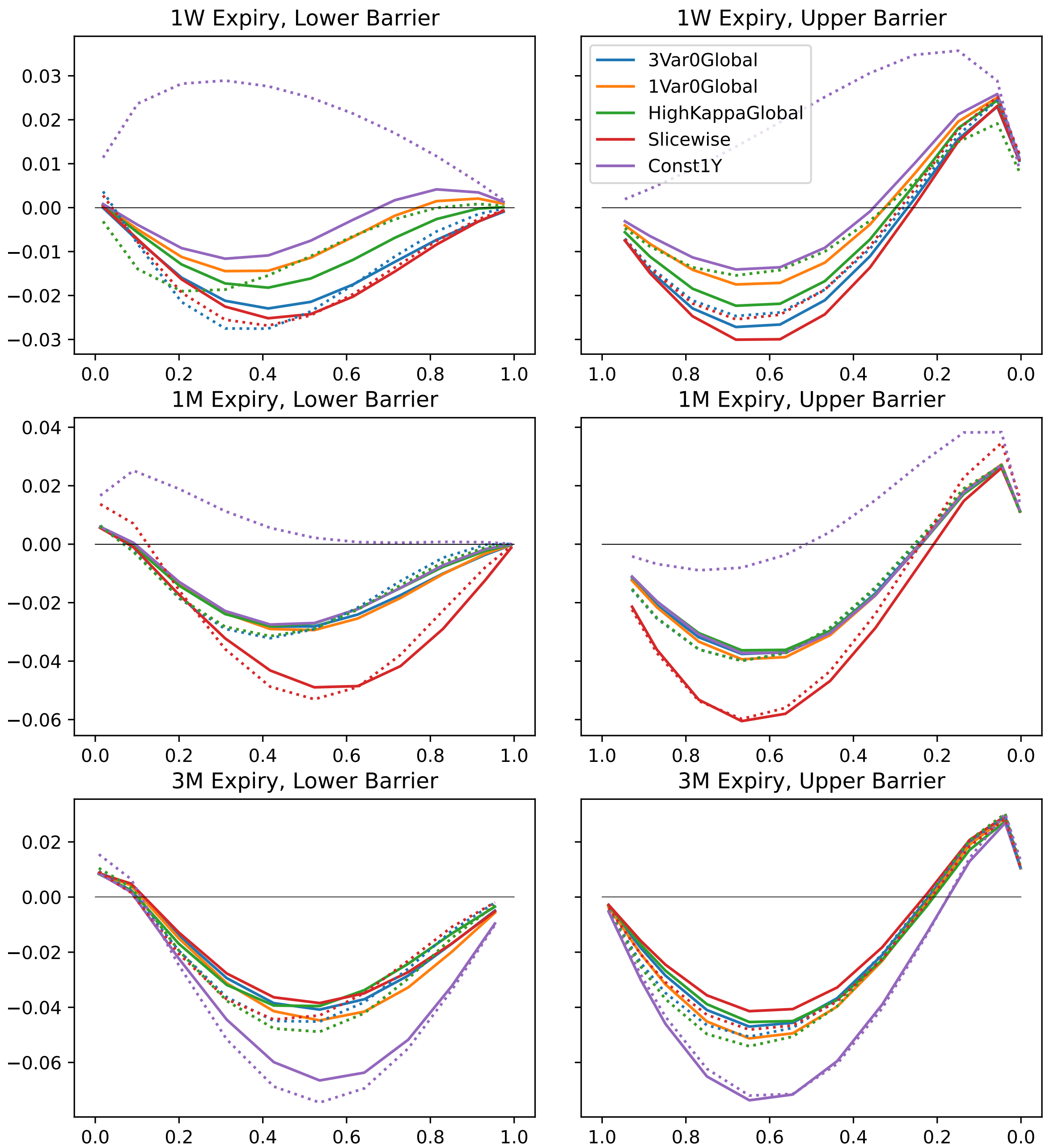


Figure 4.1: Model price differences (compared to the Black-Scholes model) for one-touch options, for some calibration alternatives for the SV model (dotted lines) and the SLV model (solid lines).

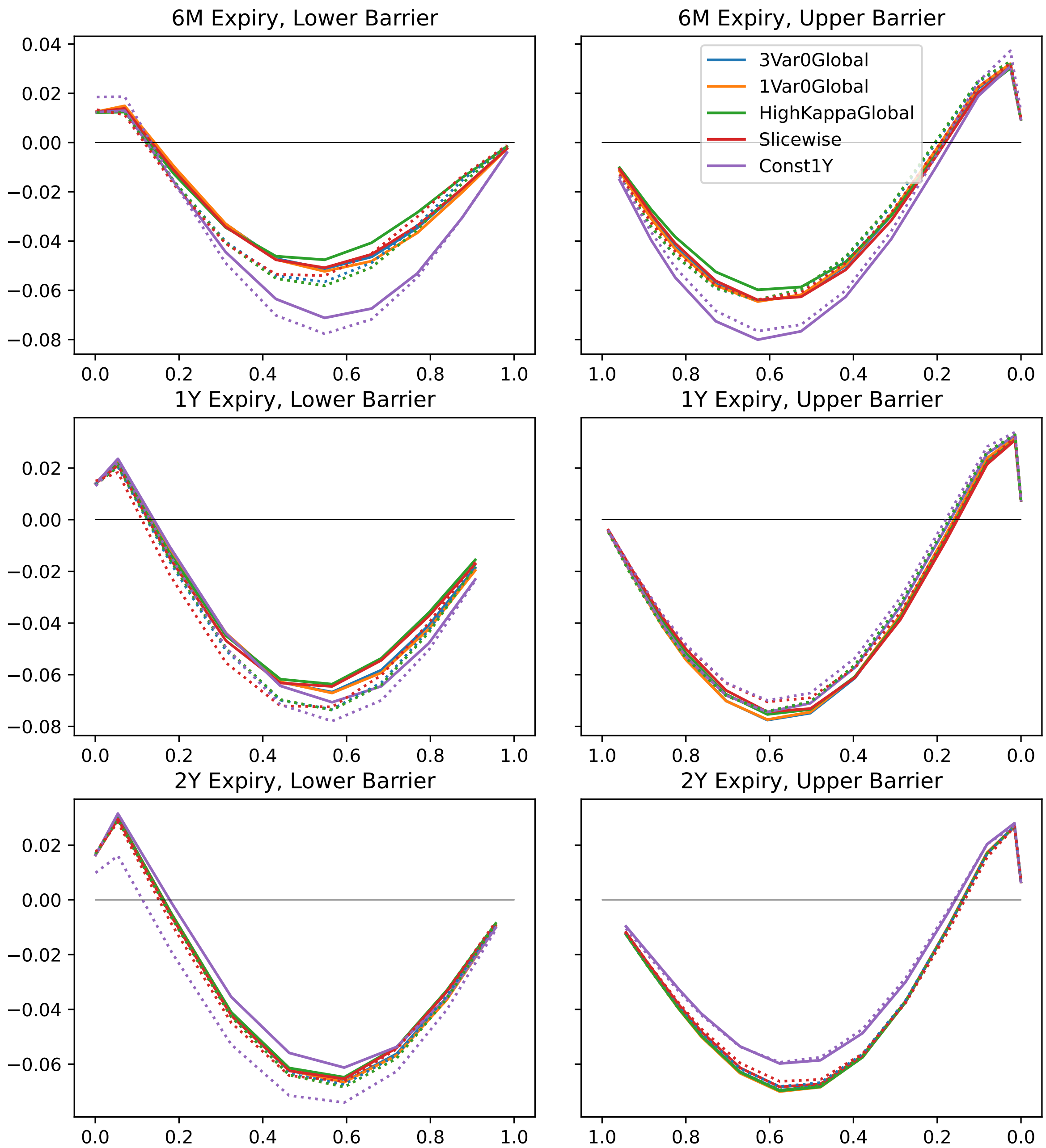


Figure 4.2: Model price differences (compared to the Black-Scholes model) for one-touch options, for some calibration alternatives for the SV model (dotted lines) and the SLV model (solid lines).

Even for the SLV models (solid lines), which fit all plain vanilla prices by construction, we observe significant differences for path-dependent claims. In particular, choosing constant SV parameters may not be sufficient.

5 Stochastic local volatility model with a mixing parameter

If the parameter ξ of the Heston model is set to zero, the variance v becomes deterministic and the model collapses to a local volatility model (LV). As traded prices of one touch contracts in FX are typically quoted somewhere between SLV and LV model prices, a standard model in FX is to mix between both.

One way to do this is to calibrate the SV model, multiply the ξ with a "mixing" factor between 0 and 1 and calibrate the leverage function $l(S_t, t)$ to fit the implied volatility surface. The mixing factor is chosen such that observed prices of one touch contracts are fitted. Alternatively, the SV model can be fitted to a smile with reduced convexity (reduced by applying the mixing factor) and the leverage function is calibrated against the original smile.

Once the mixing parameter (or a mixing parameter per tenor) has been fixed, the SLV model with mixing parameter can be used to price more general path-dependent claims.

In the following graphs we show the one touch prices (the differences to the Black-Scholes prices) under the LV model, the standard SLV model and under the mixing model with mixing factors of 0.25, 0.5, and 0.75.

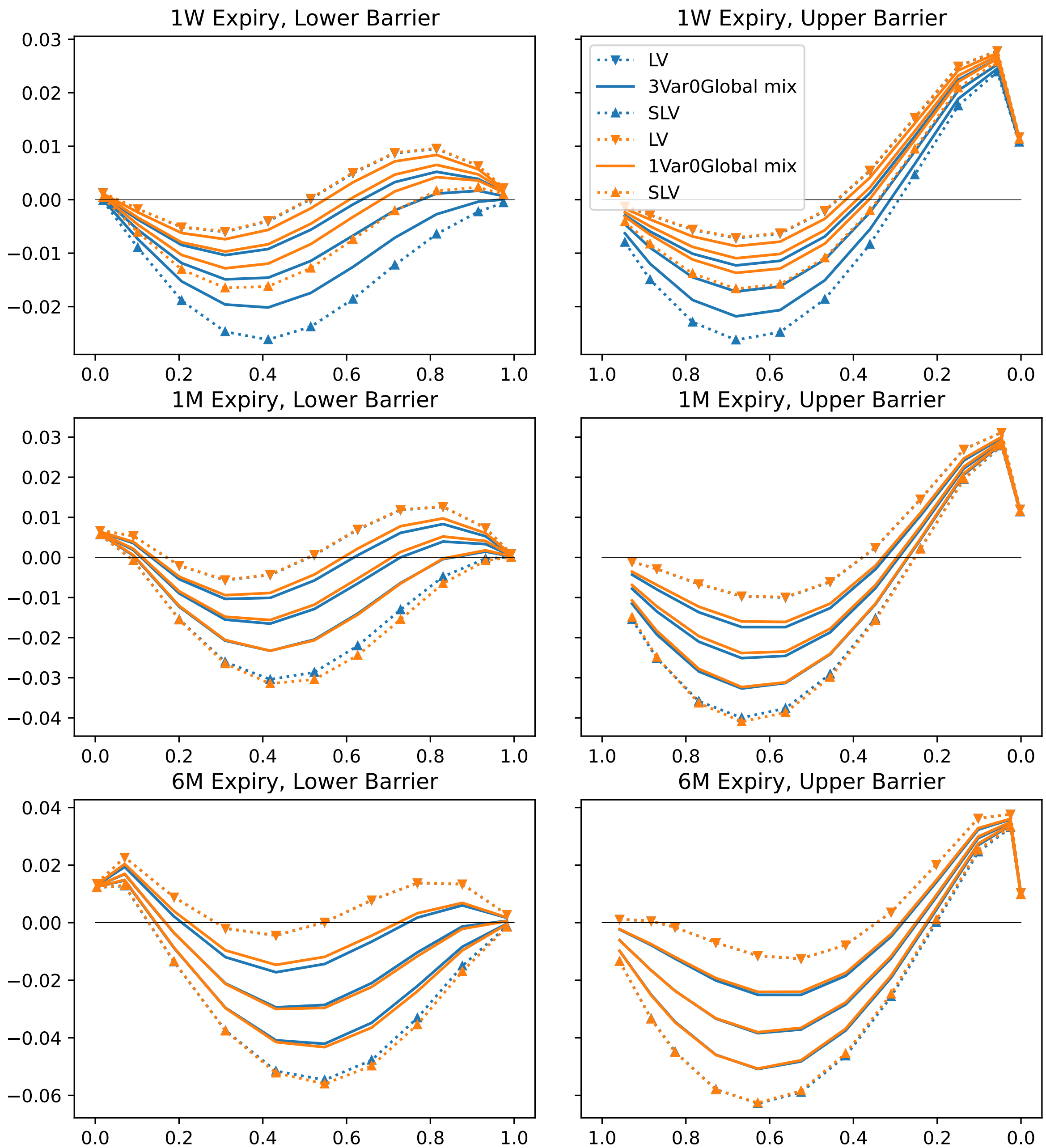


Figure 5.1: Model price differences (compared to the Black-Scholes model) for one-touch options, for SLV model with mixing factors 0 (LV), 0.25, 0.5, 0.5, 1.0 (SLV). LV coincides for both the SV calibration methods.

As we saw in Section 2, using an SV model with reasonable time dependent parameters might not produce enough variety of volatility levels in short time if started from one level of volatility. In terms of implied volatilities this yields a too flat smile for short tenors. Adding the leverage function this can be compensated for implied volatilities.

In the context of SLV with mixing parameter this is similar to already dampening down the SV model (applying a factor < 1 to ξ), as we can not fully generate the convexity with the SV model and to compensate this with a stronger local volatility component (leverage function). As we can see in the example for the 1W tenor, the model with one start variance and a mixing factor of one (SLV) coincides with the version with already diverged start variances and a mixing factor of 0.5.

With known prices of traded one touch contracts we can further determine the mixing parameters. Here we would expect the mixing parameter to be more comparable between short and longer tenors if we start the SV process with distinct levels of volatility and generate already the short term convexity by the SV model itself, rather than compensating for it with local volatility.

6 Summary

The problem of exploding volatility of variance in the Heston model when fitting short term tenors can easily be solved by using a stochastic initial volatility; even sampling from a small discrete set of initial volatilities is sufficient.

Combining this with a global optimization can avoid zigzagging volatility of variance in the Heston model.

Term-structure of the parameters is needed to fit prices of observed path-dependent claims and can be implemented successfully in a stochastic-volatility model.

References

D. Brigo, F. Mercurio, and F. Rapisarda. 2002. "Lognormal-Mixture Dynamics and Calibration to Market Volatility Smiles." *Journal of Theoretical and Applied Finance* 5(4): 427–46.

Gatheral, J. 2004. "A Parsimonious Arbitrage-Free Implied Volatility Parameterization with Application to the Valuation of Volatility Derivatives."

Heston, S. L. 1993. "A Simple New Formula for Options with Stochastic Volatility." *The Review of Financial Studies* 6(2): 327–43.

Jacquier, A., and F. Shi. 2016. "The Randomised Heston Model." *arXiv: Pricing of Securities*. https://api.semanticscholar.org/CorpusID:157416523.

Mechkov, S. 2015a. "Fast-Reversion Limit of the Heston Model." https://ssrn.com/abstract=2418631.

———. 2015b. "'Hot-Start' Initialization of the Heston Model." https://ssrn.com/abstract=2674858.

van der Zwaard, T. 2016. "The Heston Model with Term Structure: Option Pricing and Calibration." Delft University of Technology. https://resolver.tudelft.nl/uuid:eb4a8dd4-e024-48d7-9784-4bbecbebe1f1.

Wystup, U. 2019. "FX Column: Mustache to Touch – Visualizing Model Risk for Exotics." Wilmott Magazine, July, Issue 102: 10–12.

Wystup, U. 2020. "FX Column: Mixed Local Volatility Model Boosts Distribution of Exotics." Wilmott Magazine, November, Issue 110: 34–37.