INDIAN INSTITUTE OF MANAGEMENT
AHMEDABAD • INDIA

Research and Publications

Model Risk in Pricing Path-dependent Derivatives: An Illustration

Vineet Virmani

W.P. No. 2014-03-22
March 2014

INDIAN INSTITUTE OF MANAGEMENT
AHMEDABAD – 380015
INDIA

Model Risk in Pricing Path-dependent Derivatives: An Illustration

Vineet Virmani

Abstract

Model selection and model uncertainty go hand-in-hand. However, while there is uncertainty associated with the selection of any model, the context is paramount. This study is an illustration of issues surrounding model risk when pricing products whose payoff depends crucially on forward volatility. In particular, we try and quantify model risk associated with pricing of cliquet options using stochastic volatility models.

Keywords: Model risk, Option pricing, Stochastic volatility

1 Introduction

Growth in the market for path-dependent over-the-counter derivatives worldwide (BIS 2013) has meant an increasing reliance on sophisticated pricing models. While the prices of plain-vanilla options in the Black-Scholes world are essentially model-free, the same cannot be said about complex path-dependent and typically over-the-counter products like bond options, volatility derivatives and cliquet options.

The fact that the original Black-Scholes-Merton (Black and Scholes 1973, Merton 1973) argument assumes constant interest rates and volatility means that the Black-Scholes formula can no longer be used. To price a product whose payoff depends on interest rates and future unobserved volatility means one must necessarily work with models with stochastic interest rates and volatility.

If one is working with equity derivatives which are typically short maturity products, as a first approximation one could perhaps ignore randomness in interest rates. However, the constant volatility assumption is less tenable, especially if the payoff at (or before) maturity depends on future volatility of the underlying.

Given the amount of money involved in these products (BIS 2013), over time financial market professionals and researchers have come up with ways to incorporate time varying nature of volatility in pricing models.

Market professional at Goldman Sachs and erstwhile Merill Lynch first came up with "local volatility" models (Dupire 1994, Derman and Kani 1994). These models could adequately take care of information contained in option prices 'today'.

These were followed by the first generation stochastic volatility models (Hull and White 1987, Heston 1993) which addressed the limitations of the local volatility models, while still keeping providing in calibration. The second generation models followed, which were

improvements on the models of the Heston-type but with even more flexibility in incorporating market views into model parameters, while at the same time being more appealing theoretically.

Since the "true" model is only an abstraction, model selection and model uncertainty go hand-in-hand. As someone said, all models are necessarily wrong, it is only a question of how wrong! However, while there is uncertainty associated with selection of any model, the context is important. The notion of model uncertainty cannot be looked at in isolation. It only makes sense in the context of the instrument/portfolio being priced/hedged.

The objective of this note is to illustrate the nature of model risk when pricing cliquet options (often simply referred to as cliquets) using stochastic volatility models.

The plan of the article is as follows. The next section provides a very brief overview of modern stochastic volatility models. The third section describes the product (cliquet option), before briefly discussing the issue of model risk in the fourth section. The fifth and sixth sections respectively discuss the results and issues around calibration, and the seventh section concludes.

2 Overview of Stochastic Volatility Models

Use of Black-Scholes to price (and quote) options is not completely free from irony. In the Black-Scholes world, volatility is assumed to be a constant, but option traders quote that price in terms of implied volatility.

Let's make things concrete and introduce some notation at this stage. We write the option prices in the Black-Scholes world as:

$$\begin{aligned}\text{Call Option } &\equiv C(S,K,T-t,\sigma) = SN(d_1) - Ke^{-r(T-t)}N(d_2)\\ \text{Put Option } &\equiv P(S,K,T-t,\sigma) = Ke^{-r(T-t)}N(-d_2) - SN(d_1)\end{aligned}$$

$$d_1, d_2 \equiv f(S,K,r,\sigma,T-t)$$

where S is the spot, K is the strike/exercise price, $T-t$ is the time to maturity and r and σ respectively the assumed constant interest rate and volatility.

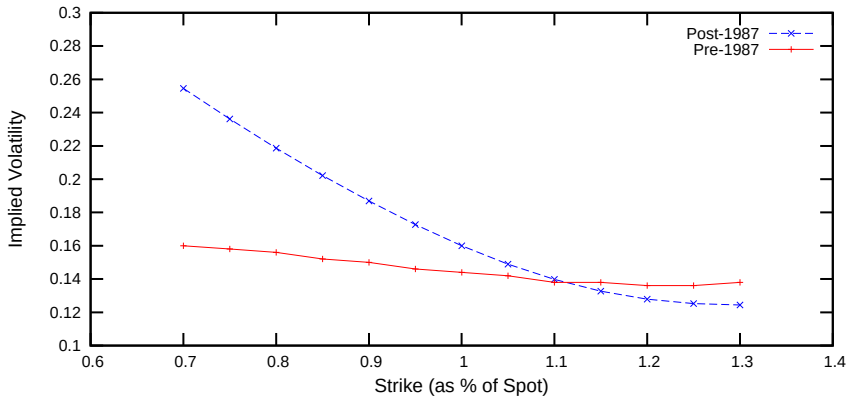
Given the option price and other inputs to the Black-Scholes model, i.e. r, $T-t$ and K, it is possible to talk about options in terms of volatility, and this $\sigma(K;C,S,r,T-t)$ is what is called as implied volatility.

This is the number quoted by option traders and not the price. It is intuitive because an option is basically a bet on the future volatility, so it makes sense to quote what is being priced. The problem is that it is not a constant. Implied volatility is known to vary with strike (the so-called phenomenon of 'volatility smile') and time to maturity (forming what is known as the 'volatility surface').

It turns out that this phenomenon of volatility smile came into being essentially after the October, 1987 crash ('Black Monday'). It was as if the 'fear' of crash got embedded in the market psyche for posterity. Figure 1 illustrates the difference in volatility smile typically observed today compared to 'back then'.

One of the reasons put options are popular (and expensive) when 'times are bad' is that they provide an insurance against the steep smile observed for low-strike options (times of high volatility or higher risk that is).

Figure 1: Smile Pre- and Post-1987 Crash (Nikkei 225)



One of the reasons put options are popular (and expensive) when 'times are bad' is that they provide an insurance against the steep smile observed for low-strike options (times of high volatility/higher risk).

The cliquet options as we'll see later also have the same property. In fact, while put option provide against today's smile, cliquet options provide insurance against its evolution.

While reviewing volatility models is not in the scope of this short note (and which can be found, for example, in Rebonato (2004)), to set the context we briefly describe different classes of stochastic models below.

2.1 Local Volatility Models

Local volatility models popularized by Derman and Kani (1994) and Dupire (1994) in mid-'90s were essentially giving a nod to the observed smile and making volatility dependent on the value of the underlying.

The original Dupire (1994) model can be written as:

$$S = \mu S dt + \sigma(S,t) dW$$

$$\frac{1}{2}\sigma^2(K,T)K^2\frac{\partial^2 C}{\partial K^2} = \frac{\partial C}{\partial t} + rK\frac{\partial C}{\partial K}$$

$$C = C(S,t,K,T,\sigma(S,t))$$

where W is the standard Brownian motion, and $\sigma(S,t)$ represents the time and spot-dependent local volatility.

The middle equation above, also known as the Dupire's equation, allows one to represent volatility in terms of the value of the underlying and time to maturity, thereby allowing

flexibility to capture the observed market smile. While this ensures that volatility would evolve in a way consistent with current implied volatility, use of such models is not without flaws.

2.1.1 Limitations of Local Volatility Models

One of the biggest concerns when using local volatility models is that the discretizing Dupire's equation to make it work is not trivial. The user is forced to use complex smoothing techniques to extract local volatility surface from a set of discrete option prices.

Given that numerical techniques have to be used, it turns out that for this particular problem, exact fit depends on *how* the implied volatility surface is smoothed. Often it is found that after interpolation, resultant local volatilities do not respect that basic put-call parity, which is a serious limitation. This often ends up requriring arbitrary fudging of the smoothing parameters/functions.

From the point of view pricing options whose value depends on dynamics of future volatility, the biggest concern is that implied forward smile from such models is observed to be flatter than the current smile. This is a contradiction, because this is like saying if I step forward one day, the model implies a smile which is flatter than the spot smile. There is no reason to expect that future put option prices would be lower than today's. The model essentially implies that.

2.2 First Generation Stochastic Volatility Models

Acknowledgment of above limitations of local volatility models led to popularity of stochastic volatility models with market professionals and researchers. While local volatility models essentially considered a deterministic volatility as a function of underlying and time, as the name suggests, stochastic volatility models started off by considering volatility as a random variable.

When compared to local volatility models, stochastic volatility models have following advantages over their counterparts:

- Such models allow one to evolve volatility independent of the stock price dynamics
- They generate smiles almost "naturally" - think of a mixture of two "high" and "low" volatility regimes
- Including a negative correlation (ρ) between the asset and the volatility allows flexibility to capture skew, which is more realistic (while also taking care of the leverage effect)
- In contrast to the local volatility models, the skew for most popular models is seen to be stable.

The most popular model in this category remains the Heston (1993) stochastic volatility model, which we briefly describe below.

2.2.1 The Heston Model

The dynamics in the Heston stochastic volatility model may be described as:

$$\begin{aligned} dS &= \mu S dt + \sqrt{\xi} S dW \\ d\xi &= \kappa(\theta - \xi)dt + \omega\sqrt{\xi} dZ \\ E[dW\, dZ] &= \rho dt \end{aligned}$$

where ξ denotes instantaneous variance, κ represents the rate of mean-reversion, θ represents the 'long-run' instantaneous variance and ω represents the 'vol-of-vol', or instantaneous variance of instantaneous volatility, and ρ the instantaneous correlation between increments of the two Brownian motions W and Z.

While the details and interpretation of the parameters may be found in Heston (1993), the model basically is a direct counterpart of the square-root Cox et al. (1985) model used in the interest rate world.

Models of the Heston-type have become popular since the late '90s. Although ultimately not very different from other one-factor affine stochastic volatility models, specifically Heston's popularity lies in the existence of a quasi-closed form solution for European options and the resulting ease of calibration.

One of the concerns when using the Heston model in practice is that given the square-root nature of the process for instantaneous variance it is not guaranteed to be positive. To ensure positivity of variance one needs the Feller condition to be satisfied, i.e. $2\kappa\bar{\nu} > \omega^2$.

Unfortunately, however, in practical calibration results, the Feller condition is often seen to be not satisfied, so if one were to use naive Euler-type discretization schemes there is a non-zero probability of the discrete process for ξ taking negative values. This is rather discomforting.

To deal with it there have been a few fixes proposed over time. The two popular ones are the "Absorption" (fixing negative variance as $\max(\xi, 0)$) and "Reflection" (fixing negative variance as $\max(\xi, -\xi)$) schemes, but both are known to result in biases (Andersen 2007).

In this study we use the method called the "Quadratic Exponential" scheme for discretization. The scheme is known to respect the martingale property under discretization and also ensures non-negativity of variance. For details please see Andersen (2007).

2.2.2 The Bates Model

A popular variant of Heston is the model proposed by Bates (1996). The Bates model is an extension of Heston in the same way as Merton (1976) is an extension to Black and Scholes (1973). Both consist of including a Poisson jump in the stochastic process for the underlying. The dynamics under Bates is:

$$\begin{aligned} dS &= \mu S dt + \sqrt{\nu} dZ_t + JS dq \\ d\nu &= -\kappa(\nu - \overline{\nu})dt + \omega\sqrt{\nu} dW_t \end{aligned}$$

Again, its popularity also lies in the existence of a quasi-closed form solution for European options (conditional on jump being uncorrelated to the two Brownians). The empirical

evidence also suggests that addition of a jump allows for a better fit to short-term options compared to Heston (Duffy and Kienitz 2009).

Working with Bates, however, is also plagued by same discretization issues as Heston. In implementation, for this also we use the Quadratic Exponential scheme for discretization.

There are other variants to these (for example, the stochastic implied volatility model of Schönbucher (1999)), but Heston and Bates capture the spirit of most models in use till the early 2000s.

2.2.3 Limitations of Heston-type Stochastic Volatility Models

A more serious problem in using stochastic volatility models is that having a separate stochastic process for variance introduces another source of randomness in the world. And if derivatives on volatility/variance are not tradable, we are in the world of incomplete markets.

This means that we can no longer use the replication argument of the Black-Scholes kind to instantaneous hedge and replicate the option, which in turn means that there would be no unique price for an option under stochastic volatility models. A unique price in stochastic volatility models exists only up to a given 'market price of risk'. If market for volatility derivatives is liquid, however, this is no more a concern.

That apart, even in traditional stochastic volatility models of the Heston-type, flexibility in evolution of forward volatility is limited. The future smile generated by these models is still not 'steep enough' for most practical purposes.

2.3 Second Generation Stochastic Volatility Models: Modeling Forward Variances

The next generation of stochastic volatility models considered by professionals were those which modeled forward variances directly. This was mostly driven by increasing demand of over-the-counter products whose payoff depended on 'forward forward volatility' (similar to 'forward forward rates' in the interest rate world).

Again, drawing analogy with the interest rate world, this is just like modeling instantaneous forward rates instead of spot rates ala Heath et al. (1992) to price interest rate derivatives.

Given a variance curve functional $G(\widetilde{z}; T-t)$, this approach goes about finding factors $\widetilde{z}_t$:

$$d\widetilde{z}_t = \mu(\widetilde{z}_t)dt + \sigma(\widetilde{z}_t)dW_t^T$$

In direct analogy with the interest rate world, where any forward-rate functional must satisfy the Heath et al. (1992) no-arbitrage condition, Buehler (2006) showed that variance curve functionals $G(\widetilde{z}; K, T-t)$ must also satisfy a similar consistency condition:

$$\partial_\tau G(\widetilde{z}_t;\tau) = \sum_{i=1}^{n} \mu_i(\widetilde{z})\partial_{z_i} G(\widetilde{z}_t;\tau) + \frac{1}{2}\sum_{i,k=1}^{n}\Big(\sum_{j=1}^{d}\sigma_i^j(\widetilde{z})\sigma_k^j(\widetilde{z})\Big)\partial_{z_i,z_k} G(\widetilde{z}_t;\tau)$$

It is interesting to note that the similarity with the theory of Heath et al. (1992) no-arbitrage condition runs deeper. Björk and Christensen (1999) showed that there does not exist any non-trivial interest rate model in the Heath et al. (1992) framework that is consistent with the Nelson and Siegel (1987) family of interest rate curves.

It turns out that the counterpart of Nelson and Siegel (1987) in the variance world is the Heston model. That is, the Heston model *does not* satisfy the Buehler (2006) consistency condition.

And the counterpart of Svensson (1994) model in the interest rate world (corresponding to which there does exist an interest rate model in the Heath et al. (1992) framework), are the "double mean reverting factor models" with the variance functionals of the same form as Svensson (1994):

$$G(\xi,\nu;\tau) = \bar{\nu} + (\xi-\bar{\nu})e^{-\kappa_1\tau} + (\nu-\bar{\nu})\frac{\kappa_1}{\kappa_1-\kappa_2}\left(e^{-\kappa_2\tau} - e^{-\kappa_1\tau}\right)$$

With this theoretical result, the Double Heston and the Double Lognormal stochastic volatility models (satisfying the Buehler (2006) condition) have become popular. We briefly describe the Double Lognormal stochastic volatility model below.

2.3.1 Double Lognormal Stochastic Volatility Model

The Double Lognormal (referred to as DLN hereafter) stochastic volatility model belongs to a class of two factor mean-reverting stochastic volatility models which have a constant elasticity of variance (CEV).

All such stochastic volatility models involve a short-term variance level that reverts to a (moving) long-term variance level at a "fast rate", with the long-term variance level reverting to a constant long-run level at a "slower rate". It was first popularized by James Gatheral of Baruch College in a 2007 Global Derivatives conference (Gatheral 2007) during his Merill Lynch days.

As the name suggests, in the DLN model the variances evolve lognormally - as against the Double Heston model, for example, where variances have a square root process for volatility (Gauthier and Possamai 2010).

In his comparison with the Double Heston model, Gatheral found that with the DLN model not only one could reproduce the market quotes better, but with DLN one could also price both options on SPX and VIX consistently (Gatheral 2008).

Note that DLN model is like a factor model for forward variances, i.e. there exist a multi-factor model for instantaneous variances that fully characterizes the evolution of forward variances over time ala Buehler (2006).

Since the quantity modeled in the DLN models is forward variances, by construction it matches variance swap prices (assuming such a market exists). Dynamics for instantaneous forward variance ξ_t^t in the DLN model are as follows:

$$d\xi = -\kappa_1(\xi-\nu)dt + \omega_1\xi dZ_1$$
$$d\nu = -\kappa_2(\nu-\bar{\nu})dt + \omega_2\nu dZ_2$$

where the stochastic process for ξ represents the "fast process", and that for ν the "slow process". Details of workings of the model for implementation are available in Appendix A.1.

3 Product: Cliquet Option

In this study we use an additive cliquet option to illustrate the sensitivity of prices of over-the-counter derivatives to the choice of the model used to price it.

3.1 Characteristics

A cliquet option is essentially an insurance on future negative returns. Let $R_n = \ln \frac{S(t_n)}{S(t_{n-1})}$ denote the return on the asset S over the interval $[t_{n-1}, t_n]$, and let U_n denote the valuation rule given by:

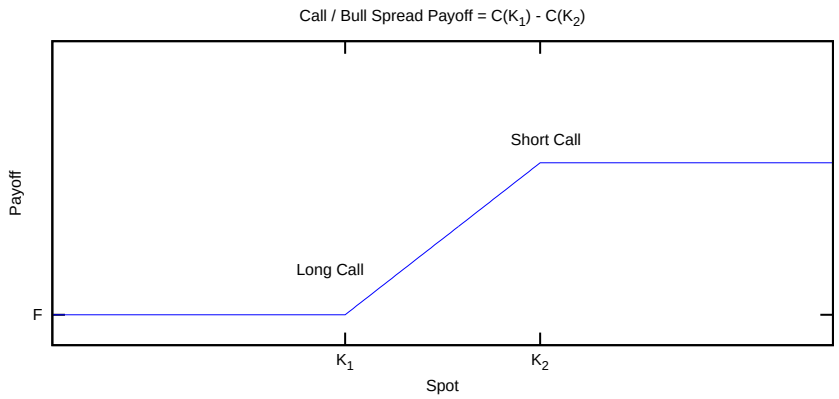
$$U_n = Min\Big(Max\big(LocalFloor, R_n\big), LocalCap\Big)$$

Given this valuation rule, cliquet options typically come in the following two avatars:

- **Additive Cliquet**: with the payoff as: $Min\Big(Max\big(GlobalFloor, \sum_{n=1}^{N} U_n\big), GlobalCap\Big)$
- **Geometric Cliquet** with the payoff as: $Min\Big(Max\big(GlobalFloor, \prod_{n=1}^{N} U_n\big), GlobalCap\Big)$

Note that each element of the sum comprising the additive cliquet option is a typical call spread (see Figure 2 below), which is a popular 'skewness trade' in the equities markets. It is essentially a portfolio of a long high strike call option and a short low strike call option.

Figure 2: Call Spread



With the intuition of a cliquet as a portfolio of call spreads, the product has the following features:

- The product is essentially going long periodically a call spread which is floored and capped. That is, another interpretation of an additive cliquet is a portfolio of call spreads on forward setting strikes.
- Global Floor/Cap ensures that consecutive negative/positive returns do not 'destroy' the buyer or the seller
- From a trading perspective, the product is designed for periods when the trader (seller) expects the performance of the asset to be between Local Floor and Cap
- The product generally trades at a premium - it offers protection against downside, while giving the benefit of potential upside
- Because it is virtually impossible to replicate, cliquets are potentially a nightmare to hedge. Typically most firms take positions on this only if 'on net basis' they have limited residual unhedged risk.

This dependence on forward setting strikes means that a cliquet's payoff depends more on 'forward volatility smile' than today's volatility smile. And since both the local volatility and the first generation stochastic volatility do not guarantee a reasonable evolution of forward smiles, pricing this is a challenge. So even if a model ensures a reasonable smile fit to today's implied volatilities, it does not guarantee a reasonable price for the cliquet.

3.2 The 'Term Sheet': Additive Cliquet Option

The specific product priced in the study is an additive cliquet option with the following 'terms':

- Cliquet type: Additive
- Underlying: XXX[1]
- Strike resetting: Monthly
- Pricing date: Sometime few years ago[2]
- Maturity: 1 year
- LocalFloor = -7%
- LocalCap = 2%
- GlobalFloor = 0
- GlobalCap $\rightarrow \infty$

The market mid-quote of the product from market participants was known to lie in the range of 2.3% - 2.6%, that is the market view at the time was that fair price of the product is around 2.45%.

[1] Permission yet to be obtained from the company from where the term sheet and other instrument details come from

[2] *ibid*

4 Model Uncertainty and Pricing of Path-dependent and Over-the-counter Derivatives

4.1 Pricing Path-dependent Over-the-counter Derivatives: Issues

The main problem in pricing path-dependent and over-the-counter derivative products lies in the fact that they are typically neither liquid nor exchange-traded.

For liquid/exchange-traded derivatives like plain vanilla European call and put options, market prices exist which reflect the fair value of the products at any point in time. So when pricing products related to today's option prices (volatility surface), the only thing that modeler has to do is to ensure that any pricing model he/she comes up with matches the market prices of these plain vanilla products. This ensures internal consistency between the model and the market prices.

However, for products such as cliquets, whose payoff depends on the *evolution* of volatility, ensuring such internal consistency is no more possible, or at least extremely difficult.

Because they are typically neither exchange-traded nor illiquid, it does not reflect the consensus market's view of the product, and if the market for forward variance (variance swaps or options on VIX for example) does not exist, their prices are also necessarily model-dependent. This introduces an additional element of risk in pricing over-the-counter products - that of model risk. To quote BIS (2009a):

"The characteristics of complex structured products...make their valuation inherently difficult due, in part, to the absence of active and liquid markets...The absence of a transparent price from a liquid market means that the valuation must rely on models or proxy-pricing methodologies, as well as on expert judgment...Moreover, calibration of the valuation methodologies is often complicated by the lack of readily available benchmarks."

As Basel recognizes the problem, it also proposes a regulation to deal with it (see BIS (2009b)):

"For complex products...banks must explicitly assess the need for valuation adjustments to reflect two forms of model risk: the model risk associated with using a possibly incorrect valuation methodology; and the risk associated with using unobservable (and possibly incorrect) calibration parameters in the valuation model."

4.2 Quantifying Model Uncertainty

While Basel recommends a "valuation adjustment to reflect ... model risk" to account for "incorrect" model and "calibration parameters" it shorts stop of suggesting ways for doing so. In the literature, one encounters the following three approaches (see Cont (2006) for a survey):

1. *Bayesian Model Averaging*: This approach exploits Bayes' theorem and calculates the probability of a model (M_j) being the true conditional on the value of parameters in

that model (θ_j) as:

$$p(M_j|\theta_j) = \frac{p(\theta_j|M_j) \times p(M_j)}{p(\theta_j)}$$

While appealing, this approach is computationally demanding as estimates of the posterior requires expensive Markov Chain Monte Carlo methods.

2. *Worst-case Approach*: This is similar to the coherent risk measures approach used in market risk management (see Artzner et al. (1999)). While it has the advantage of familiarity from market risk side, this approach does not distinguish between hedgeable vs. non-hedgeable risks and traded (liquid) vs. non-traded (illiquid) securities.

3. *Rama Cont's Approach*: Cont (2006) in his study specifically focuses on quantifying model uncertainty in the context of derivatives valuation. His approach takes into account hedging strategies in a model-free way and distinguishes explicitly between hedgeable vs. non-hedgeable risks. In doing so, he proposes the following two ways of quantifying model uncertainty:

   - *Coherent risk measure-based approach*: Given a class of models Q, for a product/payoff X (possibly neither liquid nor "replicable"), define

   $$\overline{\pi}(X) = \sup_Q E[X]$$
   $$\underline{\pi}(X) = \inf_Q E[X]$$

   Then the impact of model uncertainty on the value of the derivative X is given by $\mu(X)$:

   $$\mu(X) = \overline{\pi}(X) - \underline{\pi}(X)$$

   - *Convex risk measure-based approach*: This entails starting with a few "simpler" models (without having all the necessary features perhaps) that can be calibrated relatively easily. Then, take a larger class of models - not necessarily calibrated - but penalize each model by its error on the benchmark instruments. The idea is that

   "...easy-to-calibrate model anchors our measure of model uncertainty in the market prices while more realistic models...can be incorporated without having...heavy numerical procedures for their calibration."

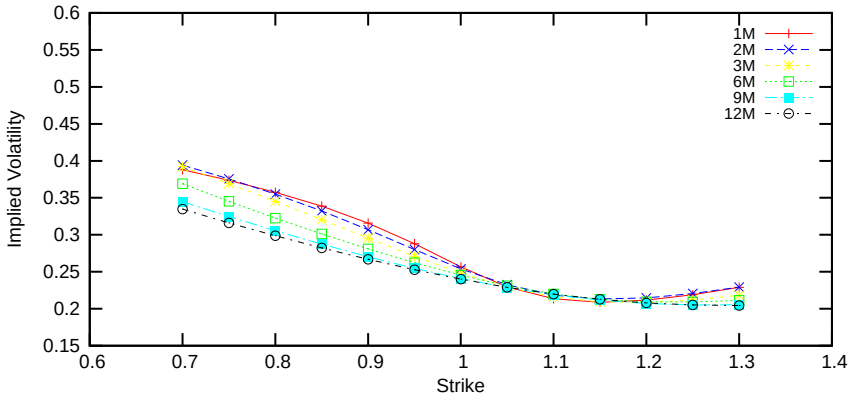
5 Results

In this section we present results for the case of additive cliquet option using the three stochastic volatility models described above, i.e. Heston, Bates and DLN. For sake of comparison at places we also present results from the Dupire/Derman-Kani also, but given that the product being priced has little dependence on today's volatility smile, we ignore any methodological considerations around these models in this note.

5.1 Spot Skew Comparison: Market Implied Volatility Smile and Smile Implied from the Models

Before presenting the results, we first present spot skew implied from the prices of plain vanilla calls and puts as on the pricing date in Figure 3 below.

Figure 3: Spot Skew for Maturities 1M - 12M



Two things are apparent. The 1M market skew is steeper than the 12M skew. It is a typical characteristic of the index options markets in 'not-so-good' times (around the time the product was to be traded, the European debt crisis was near its peak).

It is standard practice in the market to quote skew as the difference between the implied volatility at 105% and 95% strikes. While pictures are useful, skew as calculated is a concise measure of how 'steep' or 'flat' the smile is at any given point in time.

With this, Table 1 below compares the value of 0M-1M market spot skew from the different models.

While DLN seem to over-estimate it, other than that most models seem to do a fairly decent job of capturing the 1M skew[3]. (What this suggests that appropriately calibrated local as well as stochastic volatility are adequate for pricing, though hedging is another matter.)

Table 1: Spot 1M Skew ($\sigma_{95\%} - \sigma_{105\%}$) Comparison (in %)

Skew	Market	Dupire	Heston	Bates	DLN
0M - 1M Skew	5.8	5.6	5.0	5.1	6.2

Table 2 below compares the implied forward skew from the models. (The forward skew is calculated as the model implied volatility from forward starting options.)

Note that it is the implied forward skew which is more important when modeling cliquets. The table below suggest that Dupire and DLN form the two extremes within which the skew lies.

[3]Results are similar for spot skew for other maturities

Table 2: Forward 1M Skew ($\sigma_{95\%} - \sigma_{105\%}$) Comparison (in %)

Model	$1M-1M$	$2M-1M$	$3M-1M$	$6M-1M$	$11M-1M$
Dupire	3.5	3.0	3.0	1.7	1.5
Heston	3.8	3.7	3.2	3.1	3.2
Bates	4.3	4.0	3.2	3.6	3.0
DLN	5.8	5.2	5.9	5.2	5.1

It must be added that the values from Heston, Bates and DLN are conditional on the model parameters. Calibration is an important aspect of model uncertainty when using stochastic volatility models.

Before we discuss the calibration issue in detail let us take a look at the price of the cliquet implied by the four models. If calibration is done only with prevailing prices of plain vanilla calls and puts, the 'best' set of parameters (please see Appendix A.1 for parameter values used for Heston, Bates and DLN) imply that the price of the cliquet from the four models are:

Table 3: Cliquet PV Comparison

[Calibrated to Vanillas]

Model	*Cliquet PV (in %)*
Dupire	0.5903
Heston	1.6341
Bates	1.8146
DLN	1.4126
Market mid-quote	2.4500

It is clear that Dupire which matches only the current smile, and does not control evolution of the smile, performs the worst when compared to the best available market price. It stresses the fact that when pricing products like cliquets it is important to use models which specifically take into account evolution of the smile and not just the current smile.

5.1.1 Multiple Local Minima and the Role of Initial Guess

As is well known, calibrating parameter values for stochastic volatility models is a nonlinear optimization problem requiring numerical methods. This necessary entails using 'initial guesses' to start the optimization routines. In this section we look at the important issue of the role of initial guess in the eventual pricing of cliquet and the fit to vanillas.

The question that we ask here is whether the parameters that provide equally good fit to vanillas give rise to different cliquet prices (when calibrating only to vanillas).

Table 4 below shows that cliquet value varies substantially for different sets of parameters which provide similarly good fit to vanillas.

Table 4: Variation in Cliquet PV for varying initial guesses

[Calibrated to Vanillas]

Sr.	*ObjFuncVal* $\times 10^3$	*Cliquet PV (in %)*
1	0.0467	1.3031
2	0.0480	1.3105
3	0.0486	1.3387
4	0.0519	1.3412
5	0.0525	1.3272
6	0.0879	1.3773
7	0.1143	1.4066
8	0.1160	1.3370
9	0.1204	1.4126
10	0.1362	1.4533

It is clear from the analysis then that there are multiple local minima - with parameters fitting vanillas almost equally well but giving rise to a wide range of cliquet price. Given that market's view of the price for our product is available, it would also be interesting to know the impact on pricing if the models are calibrated not only to vanillas but also to the price of the cliquet. This we discuss in some detail in the next section on calibration.

6 Calibration Analysis: The Double Lognormal Model

For any stochastic volatility model to be considered useful, it must, first of all, give a good fit to vanillas. So even if it is a product like cliquet whose price depends on the *forward* skew and not the current skew, because it will be hedged by trading in vanillas it must pass the litmus test of fitting today's implied volatility surface. This is the ideal, however, and in general one looks for a reasonable fit to both vanillas and exotics.

The objective function that needs to be minimized during calibration is a weighted sum of error squared for vanillas and cliquets, with weights appropriately chosen. However, given the complicated nature of the numerical optimization problem, the output from the calibrator is as only good as the starting guess, as in general global minimum are hard to find for such problems.

In particular, we study three issues for the DLN model[4] - the extent to which initial guesses plays a role in fitting vanillas and exotics and the model risk associated with it, the 'right' value of the weight to be assigned to the cliquet price during calibration and "day-to-day" stability of the parameters. Calibration was done using the Nelder-Mead routine in the optimx package in R.

6.1 Variance Curve Input

Before we can do any of this, however, we need to get out of the way the variance curve that is required as an input to the DLN model. DLN models instantaneous forward variances,

[4]Our discussion on calibration analysis focuses on the DLN model. Concerns for Heston and Bates are similar and hence not included.

and so like Heath et al. (1992), requires the initial forward variance curve as an input to the model.

While quotes for variance swaps are not available in the market in which the cliquet is being priced, we get around the problem by using the fact that the value of a log contract is an unbiased estimate of the variance swap strike (see Demeterfi et al. (1999)). We estimate the value of the log contract across maturity on the pricing date from within the model itself (ensuring internal consistency, though not necessarily 'external validity').

6.2 Weight on Cliquet during Calibration

We begin our calibration analysis by first trying to see if we can find a way to figure out a reasonable way/rule to assign a relative weight to cliquet price during calibration.

Seeing the impact of weight on cliquet during calibration is an empirical matter. Basically, we are trying to see if there is some level/range over where the trade-off between the fit to cliquet and vanillas is optimal.

A 'U-shaped' curve of the trade-off (with the trough presenting the minimum value of the objective function), if it exists, would be ideal. Please note that at this stage we are not talking about selection of the "best" parameters - just the impact of the choice of weight on the final fit between vanillas and cliquet conditional on the choice of parameters.

To see impact of weight, we selected 6 different weights for cliquet during calibration - hoping it would be enough to provide a picture of the trend, if one exists. We selected three weights each in the 'low' (1, 2, 5), 'medium' (10, 20, 50) and 'high' (100, 200, 500) range. For evaluating the trade-off between vanilla and cliquet fit we used the the initial guess for all other parameters as in Appendix A.2.

To measure the goodness of fit with vanillas we looked at three different measures - sum of squared errors (SSE), vega-weighted sum of squared errors (VwSSE), and inverse-vega-weighted sum of squared errors (IVwSSE). While SSE gives equal weights to fit for both near-the-money options and the wings, VwSSE gives higher weight for near-the-money options and IVwSSE gives higher weight to out-of-the-money options.

The attempt is to see if the choice of measure has any bearing on the selection of final parameters. The ideal would be that all three measures won't differ too much.

To measure the goodness of fit with cliquet, we just looked at the percent deviation from the market mid-quote of 2.45%. Although we looked and studied all the measures in their own capacity, the final (minimized) objective function value thrown by the calibrator provides a neat single measure of the trade-off.

The tables below provides the ranks on all the above measures for different weights for the discrete DLN (just like the way the term 'rank' is used in other walks of life, a lower number for rank denotes a "higher" rank, that is a better fit). It is good to see that the rank to vanilla fit is invariant to choice of measure - all three SSE, VwSSE, IVwSSE agree on their ranks. Further, as expected a high rank (lower value) on vanilla on average corresponds to lower rank (higher value) on cliquet fit.

Table 5: Comparison of Goodness-of-Fit with varying weight on Cliquet during Calibration

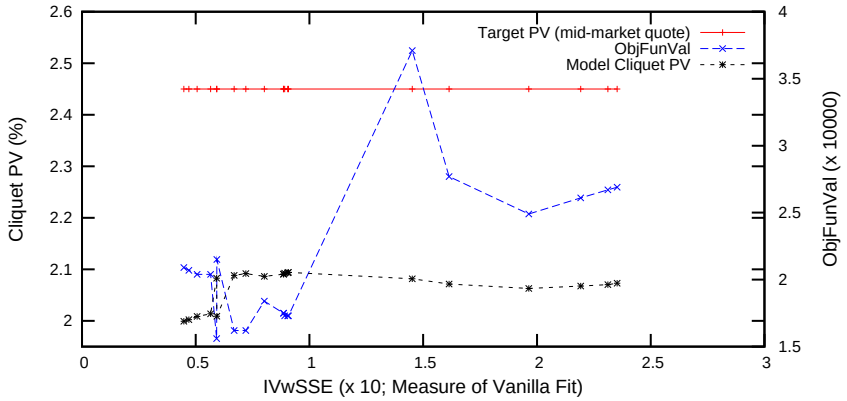
Weight	RankSSE	RankVwSSE	RankIVwSSE	RankCliquetPrice	RankObjFunVal
1	2	2	2	5	3
5	3	3	3	4	1
10	1	1	1	6	2
20	4	4	4	3	4
50	5	5	5	1	5
100	6	6	6	2	6

Given the above analysis the "right" weight on cliquet seems to depend on what is given priority, fit to vanillas or that to cliquet price.

At the risk of going out on a limb, we conjecture that the calibrator "taking" vanillas of each maturity as "one product" - one implied volatility surface if you will. With that analogy, if we are using 6 maturities - that is 6 vanilla "products" we must have a weight on cliquet closer to that number, otherwise we are over-weighting vanillas compared to cliquet and vice-versa. And if sum of ranks on cliquet fit and objective function value is any indicator, the best combination is with weight set as '5'.

The results suggest that the weight on cliquet during calibration should thus be "not too far" from number of maturities of vanillas used during calibration. In practice, of course "not too far" is an empirical matter, but as a guide, we think a factor between 1 and 2 times the number of maturities of vanillas should be ok.

Figure 4: Model Cliquet PV for different parameter sets giving "equally-good" fit to vanillas



Finally we ask the question if whether the parameters that provide equally good fit give rise to different cliquet prices, when calibrating to both vanillas and the cliquet price. Table 6 below presents the results.

Again, the cliquet price seems to vary substantially for similar (minimized) objective function value suggesting local minima.

Table 6: Variation in Cliquet PV for varying initial guesses: Double Lognormal

[Calibrated to Vanillas and Cliquet]

Sr.	*ObjFuncVal* $\times 10^3$	*Cliquet PV (in %)*
1	0.5799	1.9130
2	0.6055	1.8871
3	0.6225	2.0609
4	0.6330	2.1008
5	0.6356	1.7075
6	0.6410	1.7155
7	0.6447	1.9887
8	0.6980	2.0611
9	0.7262	1.5747
10	0.7412	1.8015

7 Conclusion: Estimate of Model Uncertainty?

We consider two cases when quantifying model risk - when benchmark (cliquet) price is not available and when it is available.

Although we have not got into details of calibration of Heston and Bates, similar concerns apply and we simply report the final results below for these two models.

7.1 Estimate of Model Uncertainty ($\mu(X)$): Benchmark Price Not Available

When the benchmark (cliquet) price is not available, by Cont's coherent risk-based measure, the estimate of model uncertainty would be:

$$\mu(X) = \overline{\pi}(X) - \underline{\pi}(X)$$

In that case Table 3 above gives cliquet value from different models. And with that $\overline{\pi}(X) = 1.8146$ (Bates) and $\underline{\pi}(X) = 1.4126$ (DLN) this gives the estimate of model uncertainty as $\mu(X) = \overline{\pi}(X) - \underline{\pi}(X) = 1.8146 - 1.4126 = 0.4020$, i.e. **almost 40 basis points**.

Add to this the minimum and maximum calibration uncertainty (range of cliquet price for "equally good" fit to vanillas across models) of 23 to 45 basis points means that total estimate of model uncertainty for pricing cliquet when benchmark price is not available would 40.2 + [23, 45] basis points, that is **in the range of 63 to 85 basis points**.

Note that when the benchmark price is not known this is likely to be an underestimate, and in that case should be thought of as the lower limit on the model-based capital reserves from the risk management point of view.

7.2 Estimate of Model Uncertainty ($\mu(X)$): Benchmark Price Available

When the benchmark (cliquet) price is not available, Cont's convex risk-based measure turns out to be more useful, and the estimate of model uncertainty would be:

$$\mu(X) = \omega_{Heston} * \mu(X)_{Heston} + \omega_{Bates} * \mu(X)_{Bates} + + \omega_{DLN} * \mu(X)_{DLN}$$

While values of ω should reflect the *apriori* "confidence" in each model, given our earlier results we may yet quantify $\mu(X)$ for different models when the benchmark (cliquet) price is available, and one can calibrate taking into account the cliquet price.

Table 7 below summarizes $\mu(X)$ along with the associated calibration error for each model:

Table 7: $\mu(X)_{\text{Model}}$ = Market mid-quote - Model price + Calibration Error

Model	*Market mid-quote - Model price*	*Calibration error*	$\mu(X)_{\text{Model}}$
Heston	2.45 - 1.64 = 0.81	0.45	1.36
Bates	2.45 - 2.08 = 0.37	0.61	0.98
DLN	2.45 - 2.06 = 0.39	0.59	0.98

Given $\mu(X)$ for the three models, and the above averaging equation, it is now up to the 'model validation' team in risk management to specify confidence in each model to come up with model reserves for risk management purposes.

7.3 Conclusion

The fact that the model reserves seem to be lower when benchmark price is not available can be misleading, because the former simply represents a lower limit in the face of uncertainty. What this suggests is that in absence of benchmark/market price of forward volatility dependent exotics, we are prone to underestimating model uncertainty even with a reasonably large class of models.

When the benchmark price is known, however, the relevant model reserve figure is obtaining by averaging across models using possibly "subjective" weights for the "confidence" in each model. But even when one knows the benchmark/market price, it is still a good idea to work with a range of appropriate models to get a reliable sense of model uncertainty.

From a risk management point of view, it would be ideal to have a ready repository of models which could be put to work or extended as the need arises. But one thing is clear, one can never have too few representative models - the more the better! The important thing is that sticking to a single measure/framework of model uncertainty is probably not a good idea.

What the above analysis also suggests is that quantifying model uncertainty is unavoidably tedious. If a structured products desk is only dealing with products whose prices depend 'nicely' on the value of the traded vanillas, quantifying model uncertainty is easy and essentially boils down to mark-to-market either directly or indirectly.

When products are highly path-dependent with limited or no knowledge of their fair price, such exercise in calibration and comparison across models is also necessary. Conservatism is a good thing when setting capital reserves against model uncertainty - especially when valuations are highly model-dependent.

What we have not looked in in this note is the impact of choice of specific discretization schemes, optimization routines and the Monte Carlo methods used to price the products. One or more of these could also end up being a big concern given the nature of the exotic derivative being priced.

What is worth exploring, however, is that given the uncertainty associated with calibration, is it *really necessary* to calibrate? And what if one is dealing with an options book, can one take advantage of netting?

To be fair, most large firms/desks typically would work with a large options book across major markets. But what happens when a large path-dependent portfolio of derivatives is being traded. In that case does diversification helps in reducing risk, or do issues of model-selection and calibration get compounded?

These are some of the more harder problems, and while they are interesting problems from the vantage point of optimization, simulation and statistics, as far as risk management is concerned, they cannot be cannot be adequately addressed outside the context of a trading desk/firm.

References

Andersen, L. B. G. (2007). Efficient Simulation of the Heston Stochastic Volatility Model. *SSRN eLibrary*.

Artzner, P., Delbaen, F., Eber, J.-M., and Heath, D. (1999). Coherent measures of risk. *Mathematical Finance*, 9(3):203–228.

Bates, D. (1996). Jump and Stochastic Volatility: Exchange Rate Processes Implict in Deutsche Mark in Options. *Review of Financial Studies*, 9:69–107.

BIS (2009a). *Enhancements to the Basel II framework*. Number 157 in Basel Committee for Banking Supervision. Bank for International Settlements.

BIS (2009b). *Revisions to the Basel II market risk framework*. Number 158 in Basel Committee for Banking Supervision. Bank for International Settlements.

BIS (2013). *OTC derivatives statistics at end-June, 2013*. Statistical release. Bank for International Settlements.

Björk, T. and Christensen, B. J. (1999). Interest rate dynamics and consistent forward rate curves. *Mathematical Finance*, 9(4):323–348.

Black, F. and Scholes, M. S. (1973). The Pricing of Options and Corporate Liabilities. *Journal of Political Economy*, 81(3):637–54.

Buehler, H. (2006). Consistent variance curve models. *Finance and Stochastics*, 10(2):178–203.

Cont, R. (2006). Model Uncertainty And Its Impact On The Pricing Of Derivative Instruments. *Mathematical Finance*, 16(3):519–547.

Cox, J. C., Ingersoll, Jonathan E, J., and Ross, S. A. (1985). A Theory of the Term Structure of Interest Rates. *Econometrica*, 53(2):385–407.

Demeterfi, K., Derman, E., Kamal, M., and Zou, J. (1999). A Guide to Volatility and Variance Swaps. *The Journal of Derivatives*, 6(4):9–32.

Derman, E. and Kani, I. (1994). The Volatility Smiles and its Implied Tree. Technical report, Goldman Sachs Quantitative Strategies Research Notes.

Duffy, D. J. and Kienitz, J. (2009). *Monte Carlo Frameworks: Building Customisable High-performance C++ Applications*. Wiley.

Dupire, B. (1994). Pricing with a Smile. *Risk*, 1(7):18–20.

Gatheral, J. (2007). Developments in Volatility Derivatives Pricing. *Presentation at the Global Derivatives Conference, Paris*.

Gatheral, J. (2008). Consistent Modeling of SPX and VIX Options. *Presentation at the Fifth World Congress of the Bachelier Finance Society*.

Gauthier, P. and Possamai, D. (2010). Efficient Simulation of the Double Heston Model. *Available at SSRN*.

Heath, D., Jarrow, R., and Morton, A. (1992). Bond Pricing and the Term Structure of Interest Rates: A New Methodology for Contingent Claims Valuation. *Econometrica*, 60(1):77–105.

Heston, S. L. (1993). A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options. *Review of Financial Studies*, 6(2):327–43.

Hull, J. C. and White, A. D. (1987). The Pricing of Options on Assets with Stochastic Volatilities. *Journal of Finance*, 42(2):281–300.

Merton, R. C. (1973). Theory of Rational Option Pricing. *Bell Journal of Economics*, 4(1):141–183.

Merton, R. C. (1976). Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics*, 3:125–144.

Nelson, C. R. and Siegel, A. F. (1987). Parsimonious Modeling of Yield Curves. *The Journal of Business*, 60(4):473–89.

Øksendal, B. (2003). *Stochastic Differential Equations: An Introduction with Applications*. Universitext. Springer, 6th edition.

Rebonato, R. (2004). *Volatility and Correlation: The Pefect Hedger and the Fox*. Wiley.

Schönbucher, P. (1999). A Market Model for Stochastic Implied Volatility. *SFB 303 Working Paper No. B - 453*.

Svensson, L. E. (1994). Estimating and interpreting forward interest rates: Sweden 1992 - 1994. Working Paper 4871, National Bureau of Economic Research.

A Appendix

A.1 Workings for Double Lognormal (DLN) Stochastic Volatility Model

Dynamics for instantaneous forward variance ξ^t in the DLN model are as follows:

$$\begin{aligned} d\xi &= -\kappa_1(\xi - \nu)dt + \omega_1 \xi dZ_1 \\ d\nu &= -\kappa_2(\nu - \overline{\nu})dt + \omega_2 \nu dZ_2 \end{aligned}$$

Scaling the forward variances in the DLN model also are scaled to today's variance curve, and this results in the following specification:

$$\begin{aligned} d\xi &= -\kappa_1(\xi - \nu)dt + \omega_1 \xi dZ_1 \\ d\nu &= -\kappa_2(\nu - 1)dt + \omega_2 \nu dZ_2 \end{aligned}$$

Before working out a solution of the above system of equations, we first develop the standard, and fairly well known (see, for example, Øksendal (2003)) result for solution of a general linear stochastic differential equation (SDE). Since this is the central result used in the solution of the above system of equations, for completeness sake we present the complete workings here.

Consider the linear SDE for $X(t)$:

$$\begin{aligned} dX &= (e(t) + f(t)X))dt + (g(t) + h(t)X)dW \\ X(0) &= X_0 \end{aligned}$$

where e, f, g and h are deterministic functions of time, and W denotes the standard Brownian motion. For notational convenience, in the following we get rid of the time dependence wherever it's obvious.

We look for solutions of the form - an ansatz if you will:

$$X(t) = X_1(t)X_2(t)$$

where,

$$\begin{aligned} dX_1 &= f(t)X_1(t)dt + h(t)X_1 dW \\ X_1(0) &= 1 \end{aligned}$$

$$\begin{aligned} dX_2 &= A(t)dt + B(t)dW \\ X_2(0) &= X_0 \end{aligned}$$

From Itô product rule we know that for any two linear SDEs driven by the same Brownian:

$$\begin{aligned} dX_1 &= a_1(t)dt + b_1(t)dW \\ dX_2 &= a_2(t)dt + b_2(t)dW \end{aligned}$$

the following holds:

$$d(X_1X_2) = X_2dX_1 + X_1dX_2 + b_1b_2dt$$

Plugging back the SDEs for X_1 and X_2 in the expression for $X(t)$, and using the Itô product rule we get:

$$\begin{aligned}
dX &= X_2dX_1 + X_1dX_2 + hBX_1dt \\
&= X_2(fX_1dt + hX_1dW) + X_1(Adt + BdW) + hBX_1dt \\
&= (f(X_1X_2) + AX_1 + hBX_1)dt + (h(X_1X_2) + BX_1)dW \\
&= (fX + AX_1 + hBX_1)dt + (hX + BX_1)dW \\
&= fXdt + hXdW + [(AX_1 + hBX_1)dt + BX_1dW]
\end{aligned}$$

Then, for our guess to make sense, we must have the last term in brackets above satisfy:

$$(AX_1 + hBX_1)dt + BX_1dW = edt + gdW$$

Equating terms for dt and dW suggests the functional forms for A and B:

$$\begin{aligned}
A &= (e - gh)(X_1)^{-1} \\
B &= g(X_1)^{-1}
\end{aligned}$$

This completes our characterization of the process $X(t)$:

$$\begin{aligned}
dX &= X_2dX_1 + X_1dX_2 + hg(X_1)^{-1}dt \\
dX_1 &= fX_1dt + hX_1dW \\
dX_2 &= (e - gh)(X_1)^{-1}dt + g(X_1)^{-1}dW
\end{aligned}$$

Having successfully parameterized the SDEs for both X_1 and X_2, it now only remains to solve them, and we are done. Since both A and B are functions of X_1, it helps that the SDE for X_1 has a simple lognormal form.

Solution for a lognormal SDE, of course, is well-known. This gives us the solution for X_1 as:

$$X_1(t) = \exp\big(\int_0^t (f(s) - \frac{1}{2}h(s)^2)ds + h(s)dW\big)$$

Conditional on A and B, solution to X_2 is trivially given by:

$$X_2(t) = X_0 + \int_0^t A(s)ds + \int_0^t B(s)dW$$

However, since A and B both depend on X_1, conditional on X_1, we can write the solution to X_2 as:

$$X_2(t) = X_0 + \int_0^t (e(s) - g(s)h(s))(X_1(s))^{-1}ds + \int_0^t g(s)(X_1(s))^{-1}dW$$

Finally, we just need to plug the expressions for X_1 and X_2 above to get a general solution for the linear SDE in X. Applying this to the case at hand is now mostly a matter of algebra. We do that now.

A.1.1 Discretization Scheme for the "Slow Process"

We start with the solution to the "slow" process ν, to which the "fast" process ξ reverts to. The SDE for ν is given by:

$$d\nu = -\kappa_2(\nu - 1)dt + \omega_2 \nu dW$$

Comparing with the general SDE solved above, we have for the "slow" SDE:

$$\begin{aligned} e &= \kappa_2 \\ f &= -\kappa_2 \\ g &= 0 \\ h &= \omega_2 \end{aligned}$$

Writing $\nu(t) = \nu_1(t)\nu_2(t)$ and comparing with results above for the general SDE in X_t, we obtain the following expressions for ν_1 and ν_2:

$$\begin{aligned} \nu_1(t) &= \exp\big(\int_0^t -(\kappa_2 + \frac{1}{2}\omega_2^2)ds + \omega_2 dW\big) \\ \nu_2(t) &= X_0 + \int_0^t \kappa_2(\nu_1(s))^{-1}ds \end{aligned}$$

These expressions don't look like that they are going to lead to an analytical solution $\nu(t)$. However, we are not interested in a solution for $\nu(t)$ as such. What is of interest is the change in $\nu(t)$ over time dt. Hopefully, the lognormal nature of the SDE will help us avoid a naïve Euler / Milstein type discretization scheme. We proceed as follows.

We write:

$$\nu(t + \Delta t) = \nu_1(t + \Delta t)\nu_2(t + \Delta t)$$

It is a matter of a few steps to show that given $\nu_1(t)$ we can write $\nu_1(t + \Delta t)$ as:

$$\nu_1(t + \Delta t) = \nu_1(t) \exp\big(\omega_2 \Delta W(t) - \kappa_2 \Delta t - \frac{1}{2}\omega_2^2 \Delta t\big)$$

Note that in the above equation, $\Delta W(t)$ is the familiar standard Brownian increment over time Δt.

Writing

$$N(\Delta t) = \exp(\omega_2 \Delta W(t) - \frac{1}{2}\omega_2^2 \Delta t)$$

we can re-write the solution for $\nu_1(t + \Delta t)$ as:

$$\nu_1(t + \Delta t) = \nu_1(t) N(\Delta t) \exp\big(-\kappa_2 \Delta t\big)$$

Similarly, given $\nu_2(t)$ as above, we can write the expression for $\nu_2(t+\Delta t)$ as:

$$\begin{aligned}\nu_2(t+\Delta t) &= \nu_0 + \int_0^{t+\Delta t} \kappa_2(\nu_1(s))^{-1} ds \\ &= \nu_0 + \int_0^{t} \kappa_2(\nu_1(s))^{-1} ds + \int_t^{t+\Delta t} \kappa_2(\nu_1(s))^{-1} ds \\ &= \nu_2(t) + \kappa_2 \int_t^{t+\Delta t} (\nu_1(s))^{-1} ds\end{aligned}$$

Given the expressions for $\nu_1(t+\Delta t)$ and $\nu_2(t+\Delta t)$, we can now expand $\nu(t+\Delta t)$ as:

$$\begin{aligned}\nu(t+\Delta t) &= \nu_1(t+\Delta t)\nu_2(t+\Delta t) \\ &= \nu_1(t)N(t)\exp(-\kappa_2\Delta t)\Big(\nu_2(t) + \kappa_2 \int_t^{t+\Delta t} (\nu_1(s))^{-1} ds\Big)\end{aligned}$$

We take a look at the integral $\int_t^{t+\Delta t} (\nu_1(s))^{-1}$ separately. Let's call it I_2. Then:

$$\begin{aligned}I_2 &= \int_t^{t+\Delta t} (\nu_1(s))^{-1} ds \\ &= \int_t^{t+\Delta t} \exp\big(-\omega_2 W(s) + (\kappa_2 + \frac{1}{2}\omega_2^2)s\big) ds \\ &= \int_t^{t+\Delta t} N(s)^{-1} \exp\big(\kappa_2 s\big) ds\end{aligned}$$

It is clear that a closed form solution for I_2 does not exist. Assuming that $N(s)^{-1} = \exp(-\omega_2 W(s) + \frac{1}{2}\omega_2^2 s)$ is linear over small time interval ds, i.e. it approximates the integral I_2 as:

$$I_2 = \frac{1}{2}\Big(N(t)^{-1} + N(t+\Delta t)^{-1}\Big) \int_t^{t+\Delta t} \exp(\kappa_2 s) ds$$

$$= \frac{1}{2}\Big(N(t)^{-1} + N(t+\Delta t)^{-1}\Big) \frac{1}{\kappa_2}\Big(\exp\big(\kappa_2(t+\Delta t)\big) - \exp\big(\kappa_2(t)\big)\Big)$$

Noting that

$$\begin{aligned}\nu_1(t) &= N(t)\exp(-\kappa_2\Delta t) \\ N(t+\Delta t) &= N(t)N(\Delta t)\end{aligned}$$

we can simplify I_2 further to get:

$$I_2 = \frac{1}{2\kappa_2}\frac{1}{\nu_1(t)}\Big(1 + N(\Delta t)^{-1}\Big)\Big(\exp\big(\kappa_2(\Delta t)\big) - 1\Big)$$

Substituting I_2 back in the expression for $\nu(t+\Delta t)$ and simplifying gives us the discretization scheme for $\nu(t)$:

$$\boxed{\nu(t+\Delta t) = \nu(t)N(\Delta t)\exp(-\kappa_2\Delta t) + \frac{1}{2}\Big(1 - \exp(-\kappa_2\Delta t)\Big)\Big(1 + N(\Delta t)\Big)}$$

A.1.2 Discretization Scheme the "Fast Process"

Given the similar nature of the SDE for the fast process, all the results for the "slow" process as above carry through. We'll skip the steps wherever it's obvious. The SDE for the "'fast"' process ξ is given by:

$$d\xi = -\kappa_1(\xi - \nu)dt + \omega_1 \nu dZ$$

Comparing with the general SDE solved above, we have for the "fast" SDE:

$$\begin{aligned} e &= \kappa_1 \nu \\ f &= -\kappa_1 \\ g &= 0 \\ h &= \omega_1 \end{aligned}$$

All the coefficients above are constants, except for $e(t)$, which here is a function of $\nu(t)$.

As earlier, we write $\xi(t) = \xi_1(t)\xi_2(t)$ and we obtain the following expressions for ξ_1 and ξ_2:

$$\begin{aligned} \xi_1(t) &= \exp\big(\int_0^t -(\kappa_1 + \frac{1}{2}\omega_1^2)ds + \omega_1 dZ\big) \\ \xi_2(t) &= \xi_0 + \int_0^t \kappa_1 \nu(s)(\xi_1(s))^{-1} ds \end{aligned}$$

Again, as earlier, we write:

$$\xi(t + \Delta t) = \xi_1(t + \Delta t)\xi_2(t + \Delta t)$$

As in the case of ν, here we have:

$$\begin{aligned} \xi_1(t + \Delta t) &= \xi_1(t) \exp\big(\omega_1 \Delta Z(t) - \kappa_1 \Delta t - \frac{1}{2}\omega_1^2 \Delta t\big) \\ M(\Delta t) &= \exp(\omega_1 \Delta Z(t) - \frac{1}{2}\omega_1^2 \Delta t) \\ \Rightarrow \xi_1(t + \Delta t) &= \xi_1(t) M(\Delta t) \exp\big(-\kappa_1 \Delta t\big) \end{aligned}$$

The main difference for ξ is in the integral for ξ_2, which now has the $e(t)$ dependent on $\nu(t)$, making it a little more complicated. We have:

$$\begin{aligned} \xi_2(t + \Delta t) &= \xi_0 + \int_0^{t+\Delta t} \kappa_1 \nu(s)(\xi_1(s))^{-1} ds \\ &= \xi_0 + \int_0^t \kappa_1 \nu(s)(\xi_1(s))^{-1} ds + \int_t^{t+\Delta t} \kappa_1 \nu(s)(\xi_1(s))^{-1} ds \\ &= \xi_2(t) + \kappa_1 \nu(s) \int_t^{t+\Delta t} (\xi_1(s))^{-1} ds \end{aligned}$$

Given the expressions for $\xi_1(t + \Delta t)$ and $\xi_2(t + \Delta t)$, we can now expand $\xi(t + \Delta t)$ as:

$$\begin{aligned}\xi(t+\Delta t) &= \xi_1(t+\Delta t)\xi_2(t+\Delta t)\\ &= \xi_1(t)M(t)\exp(-\kappa_1\Delta t)\Big(\xi_2(t)+\kappa_1\int_t^{t+\Delta t}(\nu(s)\xi_1(s))^{-1}\Big)\end{aligned}$$

As earlier, we take a look at the integral $\int_t^{t+\Delta t}\nu(s)(\xi_1(s))^{-1}ds$ separately. Let's call it I_1. Then:

$$\begin{aligned}I_1 &= \int_t^{t+\Delta t}\nu(s)(\xi_1(s))^{-1}ds\\ &= \int_t^{t+\Delta t}\nu(s)\exp\big(-\omega_1 Z(s)+(\kappa_1+\frac{1}{2}\omega_1^2)s\big)ds\\ &= \int_t^{t+\Delta t}\nu(s)M(s)^{-1}\exp\big(\kappa_1 s\big)ds\end{aligned}$$

Similar to what was done in the solution for the "slow" process, assuming that $\frac{\nu(s)}{M(s)} = \nu(s)\exp(-\omega_1 W(s)+\frac{1}{2}\omega_1^2 s)$ is linear over small time interval ds, i.e., in this case, it approximates the integral I_1 as:

$$I_1 = \frac{1}{2}\Big(\frac{\nu(t)}{M(t)}+\frac{\nu(t+\Delta t)}{M(t+\Delta t)}\Big)\int_t^{t+\Delta t}\exp(\kappa_1 s)ds$$

$$I_1 = \frac{1}{2}\Big(\frac{\nu(t)}{M(t)}+\frac{\nu(t+\Delta t)}{M(t+\Delta t)}\Big)\big(\exp\big(\kappa_1(t+\Delta t)\big)-\exp\big(\kappa_1(t)\big)\big)$$

From here on it's exactly the same as for $\nu(t)$. Noting that

$$\begin{aligned}\xi_1(t) &= M(t)\exp(-\kappa_1\Delta t)\\ M(t+\Delta t) &= M(t)M(\Delta t)\end{aligned}$$

we can simplify I_1 further to get:

$$I_2 = \frac{1}{2\kappa_1}\frac{1}{\xi_1(t)}\Big(\nu(t)+\frac{\nu(t+\Delta t)}{M(\Delta t)}\Big)\Big(\exp\big(\kappa_2(\Delta t)\big)-1\Big)$$

Substituting I_1 back in the expression for $\xi(t+\Delta t)$ and simplifying gives us the discretization scheme for $\xi(t)$:

$$\boxed{\xi(t+\Delta t) = \xi(t)M(\Delta t)\exp(-\kappa_1\Delta t)+\frac{1}{2}\Big(1-\exp(-\kappa_1\Delta t)\Big)\Big(\nu(t+\Delta t)+M(\Delta t)\nu(t)\Big)}$$

A.2 Parameters for Heston & Bates

Initial Guess: Heston & Bates

ω	κ	θ	ρ	v_0	λ	μ_J	σ_J
1.00	1.30	0.07	-0.70	0.09	0.20	-0.1	0.20

Final Estimate: Heston & Bates [Calibrated to Vanillas]

Model	ω	κ	θ	ρ	v_0	λ	μ_J	σ_J
Heston	0.875	1.477	0.072	-0.618	0.093	-	-	-
Bates	0.873	1.271	0.069	-0.615	0.093	0.237	-0.100	0.010

Final Estimate: Heston & Bates [Calibrated to Vanillas and Cliquet]

Model	ω	κ	θ	ρ	v_0	λ	μ_J	σ_J
Heston	0.875	1.477	0.072	-0.618	0.093	-	-	-
Bates	0.867	0.853	0.065	-0.617	0.099	0.738	-0.093	0.011

A.3 Parameters for Double Lognormal

Initial Guess: Double Lognormal

ω_1	κ_1	ω_2	κ_2	$\hat{\rho}$	ρ_{SX}	ρ_{SY}
7.05	13.00	1.18	0.45	0.65	-0.63	-0.63

Final Estimate: Double Lognormal [Calibrated to Vanillas]

Model	ω_1	κ_1	ω_2	κ_2	$\hat{\rho}$	ρ_{SX}	ρ_{SY}
DLN	6.99	13.02	1.17	0.39	0.63	-0.42	-0.64

Final Estimate: Double Lognormal [Calibrated to Vanillas and Cliquet]

Model	ω_1	κ_1	ω_2	κ_2	$\hat{\rho}$	ρ_{SX}	ρ_{SY}
DLN	6.95	13.05	1.00	0.46	0.62	-0.99	-0.25