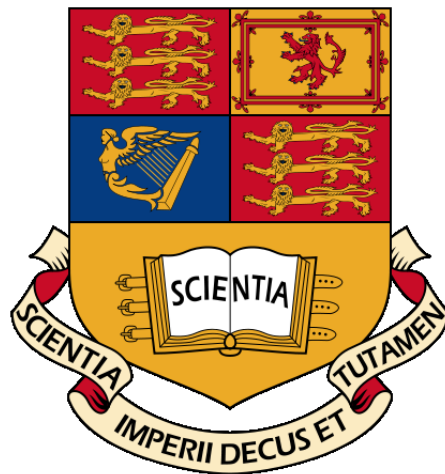


Forward Variance Dynamics: Bergomi Model And Its Applications In Pricing Cliquet Options

by

Fei Wang (CID: 01249740)

Department of Mathematics
Imperial College London
London SW7 2AZ
United Kingdom

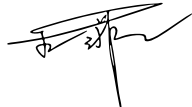


Thesis submitted as part of the requirements for the award of the
MSc in Mathematics and Finance, Imperial College London, 2016-2017

Declaration

The work contained in this thesis is my own work unless otherwise stated.

Signature and date: 09/12/2017

A handwritten signature in black ink, consisting of several overlapping horizontal strokes and a vertical line extending downwards from the center.

Acknowledgements

First of all, I would like to express my deepest gratitude to my external supervisor Dr. Mark Richardson, fund manager at Janus Henderson Investors (London), for his time, suggestions and inspirations. I particularly appreciated his kindness and dedication. His professional guidance and valuable piece of advice helped me writing this thesis.

I am also very grateful to Janus Henderson Investors, which has accepted to take in one student from the MSc in Mathematics and Finance program at the Imperial College to complete the master thesis as a summer intern. This was a great chance for me, which gave me the opportunity to rapidly get involved in a professional environment and work along with experienced practitioners. In addition, I would like to thank David Elms, head of the Multi-Strategy team, for some insightful suggestions and support.

Last but not least, I also would like to thank my academic supervisor Dr. Antonie Jacquier, who carefully and wisely supervised my thesis throughout the past few months. I highly appreciated those useful comments and discussions. It has been enjoyable to explore volatility modelling along with a leading expert, which may hopefully yield some further research. And special gratitude goes as well to Professor Damiano Brigo, for his suggestions at the start of this project, and Dr. Mikko Pakkanen, for his tutorship during my master study.

Special thanks to my family and especially my parents, for the encouragement throughout my studies. Without them it would have been impossible for me to accomplish many academic achievements and be able to continue this path.

Contents

1	Introduction	5
2	Preliminaries	7
2.1	Cliquet products	7
2.2	Variance swaps and forward variance	9
2.3	The Heston model	12
3	The Bergomi Forward Variance Model	15
3.1	Dynamics for forward variance	15
3.1.1	A N-factor model	15
3.1.2	A one-factor model	16
3.1.3	A two-factor model	17
3.2	Dynamics for asset price	18
3.3	Approximation of the smile	18
4	Numerical Simulation	20
4.1	Term structure of volatility of VS volatility and ATMF skew	20
4.2	Vanilla smile	22
4.2.1	Variance reduction	25
4.2.2	Flat VS term structure	30
4.2.3	Non-flat VS term structure	32
4.3	Approximation of the smile	33
5	Practical Implementation	35
5.1	Sufficiency based on principle component analysis	35
5.2	Calibration	36
5.2.1	Parameterisation of initial VS curve	37
5.2.2	Market Data	38
5.2.3	Calibration of Bergomi two-factor model	45
5.2.4	Calibration of Heston model	52
5.3	Pricing	53
6	Conclusion And Further Work	56
A	Expansion of the implied volatility	58
B	Practical implementation example	60

1 Introduction

Fixed Indexed Annuities (FIAs) are customized structured products, usually sold by insurance companies to retail investors. The payoff of an investment in an FIA is often guaranteed not to fall below a certain minimum level, through a global or local floor. This feature qualifies these investment products as insurance products. Such products are typically linked to the performance of one selected equity index, such as the S&P 500 (SPX) or the Euro Stoxx 50 (SX5E), and provide the retail investor capped potential return.

In recent years, there has been impressive demand for these products originating from the US, and to some extent, European insurance sectors. According to the article of Woodall [30], the US retirement market is dominated by the sale of both variable annuities (VAs) and fixed indexed annuities. FIA products, in particular, fuel the demand for strips of forward-starting call spreads, known as cliquets. As detailed in [30], in 2015 total issuance of VAs topped \$133 billion. FIAs, meanwhile, hit \$54.5 billion - a 13% increase on 2014. Especially, investment solutions that can reduce downside risk while still offering upside potential have become more popular after the turmoil in financial markets. FIA is a product embedding such features. During a period known as the accumulation phase, the investment will track an equity index, and the investor's account is credited with its upside performance up to a cap that is imposed either a monthly, semi-annual or annual basis.

A one-year FIA with a 1.5% monthly cap and 0% one-year floor, therefore, would require an insurer to buy a cliquet call spread to provide the upside potential and the cap, i.e. buying one cliquet consisting of at-the-money calls, and another consisting of selling calls at the cap level. Therefore, the net effect of an investment in an FIA is that the retail investor is short a portfolio of forward-starting call options; in turn, the insurer selling them the product is thus long this portfolio. To hedge the exposure, the insurer will typically turn to one of the investment banks, the net effect being that the long exposure passes from the insurer to the bank. In the past, banks would have been happy to carry this risk on their books, hedging certain risks in a discretionary manner with vanillas. But in the post-crisis world of punitive capital charges and restrictive risk limits, the banks are no longer as willing to retain this risk as they once were. Therefore, they have looked to offload risk direct to other counterparties, such as mutual funds and hedge funds, to step in and purchase the portfolio. The key point is that, due to the huge flow emanating from the retail sector, these portfolios of options may trade at a discount to "fair value". In general, forward-starting product are difficult to price using classical models. In particular, it is very important that the chosen model is able to capture the dynamics for the future smiles, and is consistent with the observed dynamics of volatility smiles in reality as well.

Bergomi [3] shows that cliquet-style products are highly sensitive to the shape of the future implied volatility surface and that many of the popular option pricing models impose constraints

on the dynamics of forward skew, which results in the model not being able to capture the observed dynamics correctly. This causes significant pricing errors for path-dependent or forward-starting options, in spite of the models being accurate enough to calibrate well to observed implied volatility surfaces. Motivated by these observations, a new option pricing model has been proposed by Bergomi [4], [5], [6] (see also similar work of Bühler [11] and Gatheral [15]), in which, instead of modelling instantaneous volatility, he starts by specifying the dynamics of the entire curve of forward variance. This is philosophically similar to the HJM interest rate model. The Bergomi model, which treats the forward volatility and forward skew risks accurately, has been used as a reference in [24] to show the poor volatility modeling of the Heston model, the Barndorff-Nielsen-Shephard model and a variance-gamma model with stochastic arrival. However, there is little literature specifying details of the calibration of the Bergomi model to the market. Most studies are based on numerical simulations with sets of hypothetical parameters. Indeed, most implementations of the Bergomi model utilize a flat term structure of variance swap curve, see [4], [6], [7] and [24].

The purpose of this thesis, then, is to provide a comprehensive implementation on the calibration of the Bergomi model and investigate its properties and applications for pricing cliquet-style options. Different parameterisations of initial variance swap curve will be introduced and discussed, and it will be shown that the proposed parameterisations can provide a flexible fit of the term structure of forward variances. Moreover, three different implementations of the calibration will be introduced and compared. This additional flexibility has relevance to the valuation of cliquet-style options.

The thesis is structured as follows. Some preliminary definitions and basic setting are first introduced in Section 2. The general framework of the forward variance model, and, in particular, the Bergomi two-factor model, are discussed in Section 3. The properties of the model and the relevant numerical simulations are described in Section 4. Calibration and pricing are conducted in Section 5, and Section 6 concludes.

2 Preliminaries

2.1 Cliquet products

A cliquet, with payoff being some function of a set of relative returns of the chosen underlying, is essentially a portfolio of forward-starting options. The first option (a standard vanilla option) is active immediately and expires after a predetermined period, at which point the next option activates, basing the performance on the current at-the-money level. This means that at these dates the strike price is reset at the current level of the underlying. The payoff structure makes the cliquet options particularly sensitive to the future dynamics of the implied volatility surface. Two models, which generate the same prices of European options, can lead to quite different prices for cliquet options if the dynamics of the implied volatility associated with each model is different. Wilmott [29] has shown that the sensitivity of a cliquet option to deterministic volatility is negligible in comparison with its sensitivity to volatility dynamics.

Typically, the payoff function of a cliquet-style option incorporates local or global caps and floors, minimum or maximum functions, sums and fixed coupons. And the relative returns are typically calculated on a monthly, semi-annual or annual basis. In general, there are no analytical formulas to price cliquet-style options, even in the framework of Black-Scholes, and one typically is required to resort to the use of numerical methods.

There are a wide range of investable products embedding cliquet-style features. Before introducing the three examples, we first discuss the definitions of forward volatility and forward skew. Forward volatility, $\sigma_f^{T_1, T_2}(K)$, is the implied volatility of an option with strike K and maturity T_2 which is observed at a future time T_1 . Then forward skew is intuitively defined as $\partial\sigma_f^{T_1, T_2}(K)/\partial K$. This reflects the slope of the volatility smile as a function of the strike.

Accumulator With local cap and floor, and global floor, an accumulator is very similar to an FIA. Let T , a future point in time, be the maturity date of the contract. The payoff of an accumulator is

$$\max\left(0, \sum_{i=0}^{N-1} \max(\min(r_i, \text{cap}), \text{floor})\right), \quad (2.1)$$

where

$$r_i = \frac{S_{T_{i+1}} - S_{T_i}}{S_{T_i}}, \quad 0 = T_0 < T_1 < T_2 < \dots < T_N = T.$$

The existence of these local caps and floors makes the price of the accumulator sensitive to the forward implied volatility skew. This contract can also be forward volatility sensitive but only in cases of strong forward skew.

The intuition for this behavior is similar to the intuition for skew and volatility sensitivity of a standard one month call spread. As mentioned in [24], in the framework of Black-Scholes [9],

the call spread has negligible vega. However, the absolute value of the call spread increases in the models that take the skew into account.

Reverse Cliquet As in the previous case, the investment horizon associated with reverse cliquet is also divided into a series of equally spaced periods. With the maturity T , the payoff of a reverse cliquet is

$$\max\left(0, C + \sum_{i=0}^{N-1} r_i^-\right), \quad (2.2)$$

where

$$r_i = \frac{S_{T_{i+1}} - S_{T_i}}{S_{T_i}}, \quad r_i^- = \min(r_i, 0), \quad 0 = T_0 < t_1 < \dots < T_N = T, \quad C > 0.$$

This option is called reversed cliquet because only negative returns contribute to the final payoff. The maximum payoff at the maturity is given by the coupon C and the capital is guaranteed by the existence of a global floor.

The option is both forward volatility sensitive and forward skew sensitive. The forward skew sensitivity of the reverse cliquet can be explained intuitively. If

$$\phi := C + \sum_{i=0}^{N-2} r_i^- > 0,$$

then the value of the corresponding reverse cliquet in the last period is equal to the value of a call spread option:

$$\max\left[0, \min\left(\frac{S_{T_N}}{S_{T_{N-1}}} - (1 - \phi), \phi\right)\right],$$

starting at T_{N-1} , with maturity T_N and with strikes $1 - \phi$ and 1. Therefore the value of the corresponding forward-starting call spread has an effect on the value of the whole structure.

On the other hand, in the absence of global floor, the reverse cliquet simplifies to a long coupon plus a short position on a strip of forward-starting put options. Therefore, the reverse cliquet is sensitive to the volatility of volatility. In the framework of Black-Scholes, the lower the implied volatility, the lower the value of the forward-starting puts and, hence, the higher the price of the reverse cliquet.

Napoleon This contract consists of several building blocks. The payoff of each building block, which is settled individually, is

$$\max\left(0, C + \min_{i=0, N-1} r_i\right), \quad (2.3)$$

where

$$r_i = \frac{S_{T_{i+1}} - S_{T_i}}{S_{T_i}}, \quad 0 = T_0 < T_1 < \dots < T_N, \quad C > 0.$$

This type of contract is also analyzed in Bergomi [3], [4]. Numerical simulations in Bergomi [4] show that this option is extremely forward volatility sensitive, but almost forward skew insensitive.

In Fig.1, the sensitivities of these cliquet-style options to the variance swap volatility and the difference of the implied volatilities for the strikes $0.99F_T$ and $1.01F_T$, where F_T is the forward price for maturity T , are displayed. The results are from the simulation method which will be introduced in the following sections, and based on a flat term structure of VS volatilities. The value of the flat term structure in the Bergomi model is similar to the implied volatility in Black-Scholes model. As expected, it is clear that the accumulator is not sensitive to the variance swap volatility, but very sensitive to the forward skew. And the behaviour of the napoleon is the opposite, volatility sensitive but almost skew insensitive. For reverse cliquet, it is quite sensitive to both volatility and skew. In this sense, we could interpret the reverse cliquet and napoleon as puts on volatility.

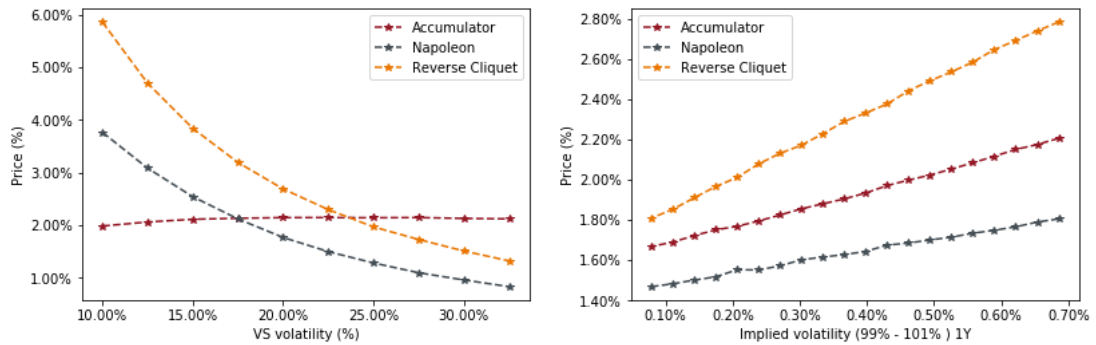


Figure 1: Sensitivity analysis of accumulator, napoleon and reverse cliquet. Left graph: sensitivity to the value of flat term structure of variance swap volatilities. Right graph: sensitivity to the difference of the implied volatilities for the strikes $0.99F_T$ and $1.01F_T$.

2.2 Variance swaps and forward variance

A variance swap (VS) with maturity T is a contract which pays out the realized variance of a financial underlying, computed as the sum of the squares of daily log-returns, in exchange for a fixed strike called the variance swap variance V_0^T . V_0^T is determined in such a way that the initial value of the contract is zero.

According to Ould Aly [27], the annualized realized variance of a stock price process $(S_t)_{t \geq 0}$ for the period $[0, T]$ with business days $0 = t_0 < \dots < t_n = T$ is usually defined as

$$RV^{0,T} := \frac{d}{n} \sum_{i=1}^n \left(\log \frac{S_{t_i}}{S_{t_{i-1}}} \right)^2.$$

The constant d denotes the number of trading days per year and is typically fixed to 252. We assume the market is arbitrage-free and prices of traded instruments are represented as conditional

expectations with respect to an equivalent pricing measure \mathbb{Q} . A standard result gives that as $\sup_{i=1,\dots,n} |t_i - t_{i-1}| \rightarrow 0$, we have

$$\sum_{i=1}^n \left(\log \frac{S_{t_i}}{S_{t_{i-1}}} \right)^2 \rightarrow \langle \log S \rangle_T \quad \text{in probability,}$$

when $(S_t)_{t \geq 0}$ is a continuous semimartingale, with $\langle \log S \rangle_t$ being its quadratic variation process. Therefore, from the perspective of expectation, there is no difference between considering realized variance or quadratic variation of returns as long as the valuation time point is one of the observation dates t_i .

The market convention of quoting a variance swap is not its variance swap rate V_0^T . Rather, the market quotes its variance swap volatility which is the strike K such that

$$\frac{1}{T} \langle \log S \rangle_T - K^2$$

has zero initial value. Therefore, we call

$$\hat{\sigma}_{VS,0}^T := \sqrt{\frac{V_0^T}{T}}$$

the variance swap volatility of the variance swap with maturity T , and the notation $\hat{\sigma}_0^T$ will be used in the rest part of the thesis for simplicity.

Let $T_1 < T_2$ be two maturities and $V_t^{T_1}, V_t^{T_2}$ the corresponding variance swap variances at time $t < T_1, T_2$. Then the forward variance swap variance $V_t^{T_1, T_2}$ is defined as

$$V_t^{T_1, T_2} := \frac{(T_2 - t)V_t^{T_2} - (T_1 - t)V_t^{T_1}}{T_2 - T_1}.$$

As in Bergomi [4], we have the following property for forward variance, which is the key point in setting the dynamics of forward variance.

Proposition 2.1. *The drift of any forward variance $V_t^{T_1, T_2}$ is zero.*

Proof. To find the drift of $V_t^{T_1, T_2}$, we first need to know the cost of entering a trade whose payoff at time $t + dt$ is linear in $V_{t+dt}^{T_1, T_2} - V_t^{T_1, T_2}$. Consider the following trading strategy with zero trading cost:

- Buy $\frac{T_2 - t}{T_2 - T_1} e^{r(T_2 - t)}$ variance swap with maturity T_2 ,
- Sell $\frac{T_1 - t}{T_2 - T_1} e^{r(T_1 - t)}$ variance swap with maturity T_1 .

Denote \hat{V}_t^T is the realized variance over the interval $[t, T]$, and then the P&L at time $t' = t + dt$ is:

$$\begin{aligned} P\&L &= \frac{T_2 - t}{T_2 - T_1} \left(\frac{\hat{V}_t^{t'}(t' - t) + V_{t'}^{T_2}(T_2 - t')}{T_2 - t} - V_t^{T_2} \right) e^{r(T_2-t)} e^{-r(T_2-t')} \\ &\quad - \frac{T_1 - t}{T_2 - T_1} \left(\frac{\hat{V}_t^{t'}(t' - t) + V_{t'}^{T_1}(T_1 - t')}{T_1 - t} - V_t^{T_1} \right) e^{r(T_1-t)} e^{-r(T_1-t')} \\ &= e^{r(t'-t)} \left(\frac{(T_2 - t')V_{t'}^{T_2} - (T_1 - t')V_{t'}^{T_1}}{T_2 - T_1} - \frac{(T_2 - t)V_t^{T_2} - (T_1 - t)V_t^{T_1}}{T_2 - T_1} \right) \\ &= (V_{t'}^{T_1, T_2} - V_t^{T_1, T_2}) e^{r(t'-t)} = (V_{t'}^{T_1, T_2} - V_t^{T_1, T_2})(1 + \mathcal{O}(dt)). \end{aligned}$$

This trading strategy generates a P&L which is linear in $V_{t+dt}^{T_1, T_2} - V_t^{T_1, T_2}$ at lowest order in dt , with zero initial cost. Therefore, the pricing drift of any forward variance is zero. \square

In order to specify the dynamics for the variance swap curve, we also need to define the value of the forward variance for the date T observed at time t as

$$\xi_t^T := V_t^{T, T}.$$

According to [27], for the continuous setting, it is given that, under \mathbb{Q} ,

$$V_t^T = \mathbb{E}_t^{\mathbb{Q}}[RV^{0, T}] = \mathbb{E}_t^{\mathbb{Q}}[\langle \log S \rangle_T].$$

We define the forward variance curve $(\xi_t^T)_{T \geq 0}$ as

$$\xi_t^T := \partial_T V_t^T, \quad T \geq t \geq 0.$$

It is easy to see that the two definitions are equivalent. Note that, if we assume that the underlying process $(S_t)_{t \geq 0}$ follows the following diffusion process:

$$dS_t = \mu_t S_t dt + \sigma_t S_t dW_t,$$

with a general stochastic volatility process $(\sigma_t)_{t \geq 0}$, then the forward variance is given by

$$\xi_t^T = \mathbb{E}_t^{\mathbb{Q}}(\sigma_T^2).$$

It can be seen as the forward instantaneous variance for date T , observed at t . In particular,

$$\xi_t^t = \sigma_t^2, \quad \text{for all } t \geq 0.$$

As mentioned by Bühler [11], the models used in practice are generally based on diffusion dynamics where forward variance curves are given as a functional of a finite-dimensional Markov-process:

$$\xi_t^T = G(T; t, Z_t),$$

where the function G and the m -dimensional Markov-process Z satisfy some consistency condition, which essentially ensures that for every fixed maturity $T > 0$, and the forward variance $(\xi_t^T)_{t \leq T}$ is a martingale.

2.3 The Heston model

Among the first generation stochastic volatility models, the Heston [22] model is perhaps the most popular one. The risk-neutral dynamics in the Heston model are as follows:

$$\begin{cases} dS_t = (r - q)S_t dt + \sqrt{V_t}S_t dW_t, \\ dV_t = -\kappa(V_t - \theta)dt + \nu\sqrt{V_t}dZ_t, \end{cases} \quad (2.4)$$

with $\langle dW, dZ \rangle_t = \rho dt$ and initial variance V_0 . All volatilities depend only on the instantaneous variance V_t , and the forward variance ξ_t^T is given by:

$$\xi_t^T = \mathbb{E}_t[V_T].$$

One of the advantages of the Heston model is its semi-closed analytical formula for the price of vanilla option. The pricing formula of a European call option with maturity T and strike K can be calculated through Fourier transform, see [2]:

$$\begin{aligned} C_0(K, T) &= \mathbb{E}(S_T - e^k)_+ \\ &= S_0 \left(\frac{1}{2} + \frac{1}{\pi} \int_0^\infty \Re \left(\frac{\Phi_T(\xi - i)}{i\xi\Phi_T(-i)} e^{-ik\xi} \right) d\xi \right) - Ke^{-rT} \left(\frac{1}{2} + \frac{1}{\pi} \int_0^\infty \Re \left(\frac{\Phi_T(\xi)}{i\xi} e^{-ik\xi} \right) d\xi \right) \\ &=: S_0\Pi_1 - Ke^{-rT}\Pi_2, \end{aligned}$$

where $\Phi_T(\xi)$ is the characteristic function of the log-stock price at time T . The characteristic function can be written in the following form:

$$\Phi_T(\xi) = \exp \left(C_T(\xi) + D_T(\xi)V_0 + i\xi \log(S_0) \right), \quad (2.5)$$

where

$$\begin{aligned} C_T(\xi) &:= ir\xi T + \frac{\kappa\theta}{\sigma^2} \left\{ (\kappa - i\rho\sigma\xi - d_T(\xi))T - 2 \log \left(\frac{1 - \gamma_T(\xi)e^{-d_T(\xi)T}}{1 - \gamma_T(\xi)} \right) \right\}, \\ D_T(\xi) &:= \frac{\kappa - i\rho\sigma\xi - d_T(\xi)}{\sigma^2} \left(\frac{1 - e^{-d_T(\xi)T}}{1 - \gamma_T(\xi)e^{-d_T(\xi)T}} \right), \\ \gamma_T(\xi) &:= \frac{\kappa - i\rho\sigma\xi - d_T(\xi)}{\kappa - i\rho\sigma\xi + d_T(\xi)}, \quad d_T(\xi) := \sqrt{(\kappa - i\rho\sigma\xi)^2 - \sigma^2(i\xi - \xi^2)}. \end{aligned}$$

The above formula is very helpful in calculating the implied volatility and quickly calibrating the model to the market.

The properties of the Heston model have been well discussed in [6], and here we only summarize the important details. Taking the conditional expectation on both sides of the second equation in (2.4) and using the notion $\bar{V}_u = \mathbb{E}_t[V_u]$ leads to:

$$d\bar{V}_T = -\kappa(\bar{V}_T - \theta)dT,$$

whose solution is

$$\bar{V}_T = \theta + e^{-\kappa(T-t)}(V_t - \theta),$$

which is equivalent to:

$$\xi_t^T = \theta + e^{-\kappa(T-t)}(\xi_t^t - \theta).$$

Differentiating the above equation, we have:

$$d\xi_t^T = \nu e^{-\kappa(T-t)} \sqrt{\xi_t^t} dZ_t. \quad (2.6)$$

Note that ξ_t^T is driftless, as it should be from Proposition 2.1. The Heston model is thus a one-factor model for forward variances. The instantaneous volatility of all forward variances ξ_t^T is proportional to the instantaneous volatility $\sqrt{\xi_t^t}$. And it is a Markov-functional model for forward variances, as ξ_t^T is a function of ξ_t^t .

However, as will be discussed in the next section, the Heston model is not a specific case of the Bergomi forward variance model. In particular, according to Bergomi [6], the one-dimension Markov representation exists only if the initial values $\xi_{t=0}^T$ of forward variances satisfy the following condition:

$$\frac{d\xi_0^T}{dT} = -\kappa(\xi_0^T - \theta)dT.$$

The Heston model is not able to generate general term structures of VS volatilities.

The parameter κ denotes the mean-reversion speed. The reciprocal of this parameter $\tau = 1/\kappa$ separates the asymptotic behaviours of volatility of volatility and at-the-money-forward (ATMF, in the sense of the strike being the forward price) skew for short and long maturities. For the Heston model, the VS volatility $\hat{\sigma}_t^T$ is given by:

$$(\hat{\sigma}_t^T)^2 = \frac{1}{T-t} \int_t^T \xi_t^\tau d\tau = \theta + \frac{1 - e^{-\kappa(T-t)}}{\kappa(T-t)} (V_t - \theta).$$

Then the dynamics of $\hat{\sigma}_t^T$ are:

$$d\left[(\hat{\sigma}_t^T)^2\right] = (\dots)dt + \frac{1 - e^{-\kappa(T-t)}}{\kappa(T-t)} \nu \sqrt{V_t} dZ_t.$$

Bergomi [6] has proposed two limiting regimes for the term structure of volatilities of volatility:

$$\begin{cases} T-t \ll 1/\kappa & \text{Vol}(\hat{\sigma}_t^T) \approx 1 - \kappa(T-t)/2, \\ T-t \gg 1/\kappa & \text{Vol}(\hat{\sigma}_t^T) \approx 1/\kappa(T-t). \end{cases} \quad (2.7)$$

Thus, for long maturities, the instantaneous volatility of $\hat{\sigma}_t^T$ decays like $1/(T-t)$, while in practice, a power-law fit is more common.

Similarly, we have the following limiting regimes for the term structure of ATMF skew in the Heston model:

$$\begin{cases} T-t \ll 1/\kappa & \mathcal{S}_T \approx \frac{\rho\nu}{4\sqrt{V_t}}, \\ T-t \gg 1/\kappa & \mathcal{S}_T \approx \frac{\rho\nu}{2\sqrt{V_0}} \frac{1}{\kappa(T-t)}, \end{cases} \quad (2.8)$$

and market skews of indices still display a power-law decay.

The above summary has highlighted some discrepancies between the dynamics of volatility generated by the Heston model and the one observed in reality. What makes the Heston model unsuitable for handling cliquet-style options is the lack of flexibility, rather than its inability to reproduce exactly the observed market.

3 The Bergomi Forward Variance Model

To capture forward volatility and forward skew risks accurately, instead of modelling instantaneous volatility, Bergomi [4] proposed a general framework for the dynamics of forward variance. Appended to this is a specification of the dynamics of the underlying consistent with that of variances. In this thesis, we will examine the particular choice of the “exponential kernel”, as proposed by Bergomi [4]. Other forms can also be considered, see [6] and [17]. And, for practical purposes, we would like to drive the dynamics of all of the ξ_t^T with a small number of factors.

3.1 Dynamics for forward variance

3.1.1 A N-factor model

In the general framework, we will use N Brownian motions and write the SDE of ξ_t^T as:

$$d\xi_t^T = \omega \alpha_\omega \xi_t^T \sum_i \omega_i e^{-\kappa_i(T-t)} dW_t^i, \quad (3.1)$$

where α_ω is a normalizing factor such that the instantaneous lognormal volatility of $\xi_t^{T=t}$ is ω , and the correlation between $(W_t^i)_{t \geq 0}$ and $(W_t^j)_{t \geq 0}$ is ρ_{ij} . Volatilities of volatilities are more natural objects than volatilities of variances. Therefore, as suggested by Bergomi [6], we introduce the lognormal volatility ν of VS volatility with vanishing maturity, which is the square root of ξ_t^t . Its instantaneous volatility is half that of ξ_t^t . We will have:

$$\begin{aligned} \omega &= 2\nu, \\ \alpha_\omega &= \frac{1}{\sqrt{\sum_{ij} \omega_i \omega_j \rho_{ij}}}. \end{aligned}$$

The solution of (3.1) is given by:

$$\xi_t^T = \xi_0^T \exp \left(\omega \sum_i \omega_i e^{-\kappa_i(T-t)} X_t^i - \frac{\omega^2}{2} \sum_{ij} \omega_i \omega_j e^{-(\kappa_i + \kappa_j)(T-t)} \mathbb{E}[X_t^i X_t^j] \right),$$

where the N driven Ornstein-Uhlenbeck (OU) processes $(X_t^i)_{t \geq 0}$ are defined by:

$$dX_t^i = -\kappa_i X_t^i dt + dW_t^i, \quad X_{t=0}^i = 0.$$

The instantaneous volatility of ξ_t^T is, from (3.1):

$$\omega(T-t) = (2\nu) \alpha_\omega \sqrt{\sum_{ij} \omega_i \omega_j \rho_{ij} e^{-(\kappa_i + \kappa_j)(T-t)}}. \quad (3.2)$$

Considering the VS volatility $\hat{\sigma}_t^T$ for maturity T :

$$(\hat{\sigma}_t^T)^2 = \frac{1}{T-t} \int_t^T \xi_\tau^T d\tau.$$

According to Itô's lemma, the dynamics of $\hat{\sigma}_t^T$ is given by:

$$d\hat{\sigma}_t^T = \nu\alpha_\omega \frac{1}{\hat{\sigma}_t^T} \sum_i \omega_i \left(\frac{1}{T-t} \int_t^T \xi_t^\tau e^{-\kappa_i(\tau-t)} d\tau \right) dW_t^i + (\dots)dt. \quad (3.3)$$

And denote ν_t^T for the instantaneous lognormal volatility of $\hat{\sigma}_t^T$:

$$\begin{cases} \nu_t^T = \nu\alpha_\omega \sqrt{\sum_{ij} \omega_i \omega_j \rho_{ij} f_i(t, T) f_j(t, T)}, \\ f_i(t, T) = \frac{\int_t^T \xi_t^\tau e^{-\kappa_i(\tau-t)} d\tau}{\int_t^T \xi_t^\tau d\tau}. \end{cases} \quad (3.4)$$

3.1.2 A one-factor model

Let us write

$$d\xi_t^T = \omega(T-t)\xi_t^T dW_t^T, \quad (3.5)$$

where $\omega(u) = \omega e^{-ku}$. Choosing an exponentially decaying volatility function is equivalent to driving the dynamics of forward variances with one OU process $(X_t)_{t \geq 0}$:

$$dX_t = -\kappa X_t dt + dW_t, \quad X_0 = 0.$$

X_t and its variance are given by:

$$X_t = \int_0^t e^{-\kappa(t-\tau)} dW_\tau, \quad \mathbb{E}[X_t^2] = \frac{1 - e^{-2\kappa t}}{2\kappa}.$$

Then the solution of SDE (3.5) reads:

$$\xi_t^T = \xi_0^T \exp \left(\omega e^{-\kappa(T-t)} X_t - \frac{\omega^2}{2} e^{-2\kappa(T-t)} \mathbb{E}[X_t^2] \right),$$

where ω is the lognormal volatility of $\xi_t^{T=t}$, a forward variance with vanishing maturity.

The slight difference between (2.6) and (3.5) leads to the structural restrictions of the Heston model and keeps it from being a particular version of the Bergomi model. However, the Bergomi one-factor model still cannot provide enough flexibility if we would like to capture the term structure of forward volatility and forward skew simultaneously. From (3.4), the instantaneous volatility of $\hat{\sigma}_t^T$ in the case of a flat term structure of VS volatilities at time t is:

$$\nu_t^T = \nu \frac{1 - e^{-\kappa(T-t)}}{\kappa(T-t)},$$

which is identical to the one obtained from the Heston model. Therefore, additional factors are needed to offer more flexibility.

3.1.3 A two-factor model

We now try with two OU processes $(X_t^1)_{t \geq 0}$ and $(X_t^2)_{t \geq 0}$. Denote their mean-reversion constants by κ_1, κ_2 , and the correlation between the Brownian motions driving $(X_t^1)_{t \geq 0}$ and $(X_t^2)_{t \geq 0}$ by ρ_{12} . We also introduce the mixing parameter $\theta \in [0, 1]$ and denote by α_θ the normalization constant such that the instantaneous lognormal volatility of $\xi_t^{T=t}$ is equal to 2ν . Then the Bergomi two-factor is given by:

$$\begin{cases} d\xi_t^T = (2\nu)\xi_t^T \alpha_\theta \left((1-\theta)e^{-\kappa_1(T-t)} dW_t^1 + \theta e^{-\kappa_2(T-t)} dW_t^2 \right), \\ \alpha_\theta = 1/\sqrt{(1-\theta)^2 + \theta^2 + 2\rho_{12}\theta(1-\theta)}. \end{cases} \quad (3.6)$$

And we introduce processes x_t^T defined as:

$$x_t^T = \alpha_\theta \left[(1-\theta)e^{-\kappa_1(T-t)} X_t^1 + \theta e^{-\kappa_2(T-t)} X_t^2 \right],$$

where $(X_t^1)_{t \geq 0}, (X_t^2)_{t \geq 0}$ are OU processes:

$$\begin{cases} dX_t^1 = -\kappa_1 X_t^1 dt + dW_t^1, & X_0^1 = 0, \\ dX_t^2 = -\kappa_2 X_t^2 dt + dW_t^2, & X_0^2 = 0. \end{cases} \quad (3.7)$$

And x_t^T is a driftless Gaussian process:

$$dx_t^T = \alpha_\theta \left[(1-\theta)e^{-\kappa_1(T-t)} dW_t^1 + \theta e^{-\kappa_2(T-t)} dW_t^2 \right],$$

whose quadratic variation is given by:

$$\begin{aligned} \langle dx^T, dx^T \rangle_t &= \eta^2(T-t)dt, \\ \eta(u) &= \alpha_\theta \sqrt{(1-\theta)^2 e^{-2\kappa_1 u} + \theta^2 e^{-2\kappa_2 u} + 2\rho_{12}\theta(1-\theta)e^{-(\kappa_1+\kappa_2)u}}. \end{aligned}$$

Then SDE (3.6) now simply reads:

$$d\xi_t^T = (2\nu)\xi_t^T dx_t^T.$$

Its solution is

$$\begin{aligned} \xi_t^T &= \xi_0^T f^T(t, x_t^T), \\ f^T(t, x) &= e^{\omega x - \frac{\omega^2}{2} \chi(t, T)}, \end{aligned} \quad (3.8)$$

where $\omega = 2\nu$ and $\chi(t, T)$ is given by:

$$\begin{aligned} \chi(t, T) &= \int_{T-t}^T \eta^2(u) du \\ &= \alpha_\theta^2 \left[(1-\theta)^2 e^{-2\kappa_1(T-t)} \frac{1 - e^{-2\kappa_1 t}}{2\kappa_1} + \theta^2 e^{-2\kappa_2(T-t)} \frac{1 - e^{-2\kappa_2 t}}{2\kappa_2} \right. \\ &\quad \left. + 2\theta(1-\theta)\rho_{12} e^{-(\kappa_1+\kappa_2)(T-t)} \frac{1 - e^{-(\kappa_1+\kappa_2)t}}{2\kappa_2} \right]. \end{aligned}$$

Note that ξ_t^T has a Markov representation as a function of x_t^T - a Gaussian process.

We also take $\kappa_1 > \kappa_2$ without loss of generality and call X^1 the short factor and X^2 the long factor. From (3.4), we have:

$$\begin{aligned} \frac{d\hat{\sigma}_t^T}{\hat{\sigma}_t^T} &= \nu\alpha_\theta \left((1-\theta) \frac{\int_t^T \xi_t^\tau e^{-\kappa_1(\tau-t)} d\tau}{\int_t^T \xi_t^\tau d\tau} dW_t^1 + \theta \frac{\int_t^T \xi_t^\tau e^{-\kappa_2(\tau-t)} d\tau}{\int_t^T \xi_t^\tau d\tau} dW_t^2 dW_t^2 \right) + (\dots)dt \\ &= \nu\alpha_\theta \left((1-\theta)A_1 dW_t^1 + \theta A_2 dW_t^2 \right) + (\dots)dt, \end{aligned} \quad (3.9)$$

with A_i give by

$$A_i = \frac{\int_t^T \xi_t^\tau e^{-\kappa_i(\tau-t)} d\tau}{\int_t^T \xi_t^\tau d\tau}.$$

The instantaneous volatility of a VS volatility ν_t^T is given by:

$$\nu_t^T = \nu\alpha_\theta \sqrt{(1-\theta)^2 A_1^2 + \theta^2 A_2^2 + 2\rho_{12}\theta(1-\theta)A_1 A_2}. \quad (3.10)$$

3.2 Dynamics for asset price

We would write the following lognormal dynamics on the underlying:

$$dS_t = (r - q)S_t dt + \sqrt{\xi_t^S} S_t dW_t^S, \quad (3.11)$$

where W_t^S is correlated with the Brownian motions in the dynamics of forward variance. Particularly, for a two-factor model, we can denote ρ_{12} as the correlation between W^1 and W^2 , ρ_1, ρ_2 as the correlations between W^S and W^1, W^2 respectively. This yields a stochastic volatility model which has two factors and which can be calibrated to the term-structure of VS volatilities. Also, in such a model, the level of forward skew is determined jointly by $\rho_1, \rho_2, \rho_{12}, \nu, \kappa_1, \kappa_2, \theta$.

3.3 Approximation of the smile

For the Bergomi model, the pricing equation for European options is not analytically solvable and we have to resort to Monte Carlo simulations. However, with noisy simulation results and huge computation cost, it would be a disaster if we utilize Monte Carlo to calibrate our models. To solve this problem, Bergomi and Guyon [7] has derived an approximation of the smile produced by the forward variance model at second order in the volatility of volatility. They introduce a scaling factor ϵ for the volatilities of forward variances and derive that at second order in ϵ , the implied volatility for maturity T and strike K are exactly quadratic in log-moneyness:

$$\hat{\sigma}(K, T) = \hat{\sigma}(F_T, T) + \mathcal{S}_T \ln \left(\frac{K}{F_T} \right) + \frac{\mathcal{C}_T}{2} \ln^2 \left(\frac{K}{F_T} \right) + \mathcal{O}(\epsilon^3). \quad (3.12)$$

The ATMF volatility $\hat{\sigma}(F_T, T)$, the ATMF skew \mathcal{S}_T and curvature \mathcal{C}_T are given by:

$$\hat{\sigma}(F_T, T) = \hat{\sigma}^T \left[1 + \frac{\epsilon}{4Q} C^{x\xi} + \frac{\epsilon^2}{32Q^3} (12(C^{x\xi})^2 - Q(Q+4)C^{\xi\xi} + 4Q(Q-4)C^\mu) \right], \quad (3.13a)$$

$$\mathcal{S}_T = \hat{\sigma}^T \left[\frac{\epsilon}{2Q^2} C^{x\xi} + \frac{\epsilon^2}{8Q^3} (4C^\mu Q - 3(C_{x\xi})^2) \right], \quad (3.13b)$$

$$\mathcal{C}_T = \hat{\sigma}^T \frac{\epsilon^2}{8Q^4} (4C^\mu Q + C^{\xi\xi} Q - 6(C^{x\xi})^2), \quad (3.13c)$$

where $Q = \int_0^T \xi_0^s ds$ and $\hat{\sigma}^T = \sqrt{\frac{Q}{T}}$, the VS volatility for maturity T , and $C^{x\xi}, C^{\xi\xi}, C^\mu$ summarize the joint spot/variance dynamics of the model at hand. $C^{x\xi}$ and $C^{\xi\xi}$ are integrals of the spot/variance and variance/variance covariance functions evaluated on the initial variance curve, and C^μ involves an extra degree of model-dependence as it depends on the derivative of $C^{x\xi}$ with respect to ξ . The details of $C^{x\xi}, C^{\xi\xi}, C^\mu$ are showed in Appendix A. And when we use these formulas, we will set $\epsilon = 1$.

It is clear from (3.13a) that ATMF implied volatility is the variance swap volatility plus a spread. At first order the spread is $C^{x\xi}/4Q$. Typically, for the equity market, $C^{x\xi} < 0$, and, hence, ATMF implied volatility lies below the variance swap volatility. When the correlation between the underlying and the variance is zero, $C^{x\xi} = C^\mu = 0$, we have

$$\hat{\sigma}(F_T, T) = \hat{\sigma}^T \left[1 - \frac{\epsilon^2}{32Q^3} (Q(Q+4)C^{\xi\xi}) \right].$$

The ATMF implied volatility lies again below variance swap volatility. And the higher the value of variance swap volatility, the smaller the ATMF implied volatility.

At order one in ϵ , from (3.13b), the ATMF skew is given by

$$\mathcal{S}_T = \hat{\sigma}^T \frac{C^{x\xi}}{2(\hat{\sigma}^T T)^2}, \quad (3.14)$$

where we have set $\epsilon = 1$. Whenever spot and variances are uncorrelated, \mathcal{S}_T vanishes both at order ϵ and ϵ^2 , and at all orders, as it should, since it is a well-known result that the smile is symmetric in log-moneyness for uncorrelated spot and variances.

Bergomi and Guyon conclude that in the case of the Bergomi two-factor model, there are good agreements of the order one expression for the ATMF skew, and of the order two expression for the ATMF volatility, for values of the volatility of short-dated variance (around $\nu = 200\%$) that are typical of implied levels of equity indices. The accuracy of the expansion will be verified in the following section.

4 Numerical Simulation

In this section, several numerical simulations have been implemented to test the flexibility of a two-factor model and the accuracy of expansion in (3.12). For simplicity, a flat term structure of VS volatilities has been used, if not specified particularly. Additionally, some techniques for Monte Carlo simulations are also discussed.

4.1 Term structure of volatility of VS volatility and ATMF skew

In the case of a flat term structure of VS volatilities, ξ_t^τ does not depend on τ . The integral in (3.4) and (3.10) can be evaluated analytically and for a two-factor model, we get the following simple formula for the instantaneous volatility of $\hat{\sigma}_t^T$:

$$\nu_t^T = \nu\alpha_\theta \sqrt{(1-\theta)^2 A_1^2 + \theta^2 A_2^2 + 2\rho_{12}\theta(1-\theta)A_1 A_2},$$

with A_i given by:

$$A_i = \frac{1 - e^{-\kappa_i(T-t)}}{\kappa_i(T-t)}.$$

Among the parameters of the two-factor model, the subset $\nu, \theta, \kappa_1, \kappa_2, \rho_{12}$ determines the dynamics of the VS volatilities in the model. Once these parameters are set, the dynamics of VS volatilities is set. Then we can select the additional parameters ρ_1 and ρ_2 to generate the desired spot/volatility dynamics and the desired vanilla smile.

The accuracy of expression (3.13b) for the ATMF skew \mathcal{S}_T , is excellent already at order one in ϵ , as highlighted in [7]. At this order, the formula simplifies to (3.14). For a Bergomi two-factor model, we have

$$\mathcal{S}_T^{\text{order 1}} = \frac{\nu\alpha_\theta}{(\hat{\sigma}^T)^3 T^2} \int_0^T dt \sqrt{\xi_0^t} \int_t^T du \xi_0^u \left[(1-\theta)\rho_1 e^{-\kappa_1(u-t)} + \theta\rho_2 e^{-\kappa_2(u-t)} \right] \quad (4.1)$$

where $\hat{\sigma}^T = \sqrt{\frac{1}{T} \int_0^T \xi_0^t dt}$. Particularly, in the case a flat term structure of forward variances/Vs volatilities, the double integrals can be evaluated analytically to obtain:

$$\mathcal{S}_T^{\text{order 1}} = \nu\alpha_\theta \left[(1-\theta)\rho_1 \frac{\kappa_1 T - (1 - e^{-\kappa_1 T})}{(\kappa_1 T)^2} + \theta\rho_2 \frac{\kappa_2 T - (1 - e^{-\kappa_2 T})}{(\kappa_2 T)^2} \right]. \quad (4.2)$$

Normally, for equity indexes, volatilities of VS volatilities usually display a power-law dependence on maturity, with an exponent that typically lies between 0.3 and 0.6. For the numerical simulations below, we use the same parameters from Bergomi [6].

We will use the following time-homogeneous benchmark form for ν_t^T :

$$\nu_t^{T,B} = \sigma_0 \left(\frac{\tau_0}{T-t} \right)^\alpha \quad (4.3)$$

where τ_0 is a reference maturity and σ_0 is the volatility of $\hat{\sigma}_t^{t+\tau_0}$. Typically, we will take $\alpha = 0.4$, $\tau_0 = 3$ months and $\sigma_0 = 100\%$. Table 1 displays the parameters used to match $\nu_t^{T,B}$ in (4.3) with the above setting.

	ν	θ	κ_1	κ_2	ρ_{12}
Set I	150%	0.312	2.63	0.42	-70%
Set II	174%	0.245	5.35	0.28	0%
Set III	186%	0.230	7.54	0.24	70%

Table 1: Three sets of parameters used in forward variance process

Fig.2 displays the term structure of instantaneous volatilities at $t = 0$ of VS volatilities ν_t^T as a function of T generated by the benchmark form (4.3) as well as the two-factor model for a flat term structure of VS volatilities. The three set of parameters are differentiated by the value of the correlation between processes $(X_t^1)_{t \geq 0}$ and $(X_t^2)_{t \geq 0}$, We have used $\rho_{12} = -70\%, 0, 70\%$ and have selected the remaining parameters $\nu, \theta, \kappa_1, \kappa_2$ so as to best match our benchmark (4.3) for maturities from one month to 5 years.

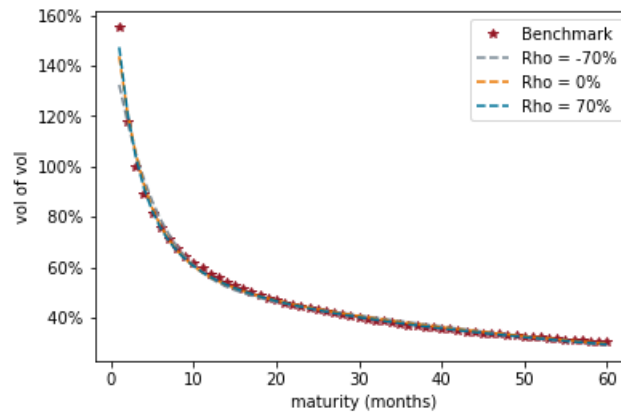


Figure 2: Term structure of instantaneous volatilities of VS volatilities generated by the benchmark function as well as the two-factor model, with the different sets of parameters listed in Table 1.

It is clear that the two-factor model is able to capture a power-law dependence for volatilities of volatilities over a wide range of maturities. Moreover, for a given α , many different sets of parameters exist that provide an equally acceptable fit to our benchmark.

The next question is how should we choose ρ_1, ρ_2 ? Once other parameters are set, ρ_1, ρ_2 will determine both the vanilla smile and future smiles. It is necessary that the two-factor model be at least able to generate smiles that are comparable to historically observed smiles, in particular with respect to the term structure of the ATMF skew. This is especially important when pricing cliquet-style options. Typically, equity index smiles display a term structure of the ATMF skew that is well approximated by a power law with an exponent usually around 0.5.

As indicated by Bergomi [6], the difference of the implied volatilities for strikes $0.99F_T$ and $1.01F_T$, where F_T is the forward price for maturity T , is approximately equal to $-0.02\mathcal{S}_T$ and

the order-one expansion is remarkably accurate. We therefore measure the ATMF skew as the difference of implied volatilities of implied volatilities for strikes $0.99F_T$ and $1.01F_T$ and calculate the skew by expression (4.2) for $\mathcal{S}_T^{\text{order } 1}$.

ν	θ	κ_1	κ_2	ρ_{12}	ρ_1	ρ_2
174%	0.245	5.35	0.28	0.0%	-75.9%	-48.7%

Table 2: Values of parameters of the two-factor model

The values for ρ_1, ρ_2 in Table 2 are such that they generate a term structure for the ATMF skew that is approximately a power law with exponent 0.5, with $\hat{\sigma}_{0.99F_T} - \hat{\sigma}_{1.01F_T} = 0.6\%$ for $T = 1$ year. As illustrated in Fig.3, the Bergomi two-factor model can generate the desired term structure of the ATMF skew.

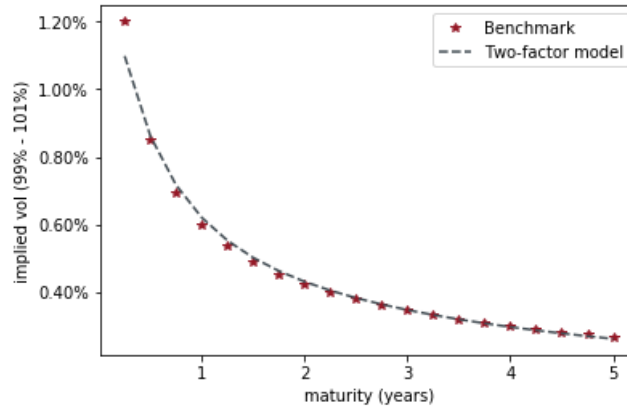


Figure 3: Term structure of ATMF skew measured as the difference of implied volatilities for strikes $0.99F_T$ and $1.01F_T$, with the parameters in Table 2. A flat VS volatility at 20% has been used.

We could as well have chosen other values for ρ_1, ρ_2 such that the ATMF skew in the two-factor model generates a power-law dependence with a different exponent. However, the freedom is limited. Indeed, the other parameters in the model are already set and the triplet $\rho_{12}, \rho_1, \rho_2$ must constitute a valid correlation matrix. This is the case if ρ_2 is defined as:

$$\rho_2 = \rho_{12}\rho_1 + \chi\sqrt{1 - \rho_{12}^2}\sqrt{1 - \rho_1^2}, \quad (4.4)$$

where $\chi \in [-1, 1]$. The above restriction will be used in both Monte Carlo simulations and the algorithm of calibration.

4.2 Vanilla smile

As mentioned before, there is no analytical formula for the price of vanilla options with the Bergomi model, and, hence, we have to implement a Monte Carlo simulation. For the Bergomi two-factor

model with (3.6) and (3.11), the Euler schemes for the two OU-processes (3.7) and the underlying process (3.11) are given below:

$$X_{t+\Delta t}^1 - X_t^1 = -\kappa_1 X_t^1 + \sqrt{\Delta t} X, \quad (4.5a)$$

$$X_{t+\Delta t}^2 - X_t^2 = -\kappa_2 X_t^2 + \sqrt{\Delta t} Y, \quad (4.5b)$$

$$\log(S_{t+\Delta t}) - \log(S_t) = \left(r - q - \frac{1}{2} \xi_t^t \right) \Delta t + \sqrt{\xi_t^t \Delta t} Z. \quad (4.5c)$$

where X, Y, Z are Brownian motions with mean 0, variance 1 and the following correlation structure:

$$\begin{aligned} X &= W^1, \\ Y &= \rho_{12} W^1 + \sqrt{1 - \rho_{12}^2} W^2, \\ Z &= \frac{\rho_1 - \rho_{12} \rho_2}{1 - \rho_{12}^2} W^1 + \frac{\rho_2 - \rho_{12} \rho_1}{1 - \rho_{12}^2} W^2 + \sqrt{1 - \lambda^2} W^3, \end{aligned} \quad (4.6)$$

and

$$\lambda = \sqrt{\frac{\rho_1^2 + \rho_2^2 - 2\rho_{12}\rho_1\rho_2}{1 - \rho_{12}^2}},$$

where W^1, W^2, W^3 are three independent standard Brownian motions. The detailed procedure for the pricing of options using Monte Carlo simulations is demonstrated in Algorithm 1.

Algorithm 1 Basic Monte Carlo algorithm

```

1: procedure OPTION PRICE( $N, M$ )                                ▷ Path number N, time step M
2:    $sum \leftarrow 0$ 
3:   for  $i = 1$  to N do
4:     generate  $W^1, W^2, W^3$  and calculate  $X, Y, Z$                 ▷ With (4.6)
5:     for  $j = 1$  to M do
6:       simulate paths of  $X_t^1, X_t^2$                             ▷ With (4.5a), (4.5b)
7:       set path of  $\xi_t^t$                                        ▷ With (3.8)
8:       simulate path of  $S_t$                                     ▷ With (4.5c)
9:     end for
10:     $sum \leftarrow sum + F(S_T)$                                 ▷ Payoff  $F(\cdot)$ 
11:  end for
12:  return  $sum/N$ 
13: end procedure

```

Denote \hat{C}_0 as an estimate of the true value of the option C_0 , but with an error due to the fact that it is an average of randomly generated samples and so is itself random. A measure of the error is the standard error

$$Se(\hat{C}_0) = \frac{Sd(\hat{C}_0)}{\sqrt{N}},$$

where the standard deviation $\text{Sd}(\hat{C}_0)$ of \hat{C}_0 is

$$\text{Sd}(\hat{C}_0) = \sqrt{\frac{1}{N-1} \sum_{j=1}^N (C_{0,j} - \hat{C}_0)^2}.$$

Another way to measure the quality of the estimator is the half-width, HW, of the confidence interval. For a fixed α , we have

$$\text{HW} = z_{1-\alpha/2} \text{Se}(\hat{C}_0).$$

The parameters used for the Bergomi two-factor model is the same as the ones used in last section, as listed in Table 2. We simulate the price of a vanilla ATM European call option with the parameters in Table 3.

S_0	K	T	r	q
50.0	50.0	1.0	0.5%	0.0%

Table 3: Values of parameters for the option

In order to get an acceptably accurate estimate of the option price, a very large number of simulations has to be performed, typically in the order of millions ($M > 1,000,000$). But the cost of computation (running time) is very large, and the speed of convergence for Monte Carlo simulation is slow, proportional to $1/\sqrt{N}$, where N is the number of simulation path. Thus, cutting the error in half requires increasing the number of points by a factor of four, and adding one decimal place of precision requires 100 times as many points. The results of standard Monte Carlo simulations are displayed in Table 4 .

Path Num	Time Step	Price	Std Error	95% HW
500,000	100	3.60	0.58%	0.63%
1,000,000	100	3.59	0.42%	0.44%
2,000,000	100	3.59	0.29%	0.31%

Table 4: Values of the option with basic Monte Carlo simulations

This problem can be ameliorated to some extent using some “variance reduction” techniques.

Remark 4.1. The speed of Python code can often be increased greatly by vectorizing mathematical expressions that applied to NumPy arrays rather than using loops. The results can be observed from the following simulations in Table 5. The first one is based on code using only loops (two for-loop: time discretization and path simulation), while the second one is based on vectorization. The code with vectorization is far much faster, but still not fast enough yet to be used in the calibration.

Remark 4.2. Another way to increase the simulation speed is to use Cython, which is an optimizing static compiler for the Python programming language. Cython provides the combined power of Python and C to let you interact efficiently with large data sets, e.g. using multi-dimensional NumPy arrays. The results can also be observed from the following simulations in Table 5. The code with Cython performs slightly better than the code with vectorization.

Code Type	Path Num	Time Step	Price	Std Error	95% HW	Run Time
For-loops	125,000	100	3.58	1.17%	1.25%	168.14s
Vectorization	125,000	100	3.60	1.18%	1.25%	4.39s
Cython	125,000	100	3.59	1.17%	1.25%	3.84s

Table 5: Run time of different code types

4.2.1 Variance reduction

The standard Monte Carlo of averaging vanilla option payoff over all simulated paths produces price estimations that are in practice too noisy to use. In this subsection, we present several more efficient techniques.

Antithetic variates: Suppose we would like to estimate $E[Y]$, and that we have generated two samples, Y_1 and Y_2 . Then an unbiased estimate is given by

$$\frac{Y_1 + Y_2}{2},$$

with variance

$$\frac{\text{Var}(Y_1) + \text{Var}(Y_2) + 2\text{Cov}(Y_1, Y_2)}{4}.$$

We could reduce the variance if we could arrange it so that $\text{Cov}(Y_1, Y_2) < 0$. The method of antithetic variates is based on this idea. Set $Y_i = h(W_i)$, where $W_i = (W_1^{(i)}, \dots, W_m^{(i)})$ is a vector of independently identically distributed standard normal variables. We now also set $\tilde{Y}_i = h(-W_i)$. Note that $E[Y_i] = E[\tilde{Y}_i] = E[Y]$ so that in particular, if

$$Z_i = \frac{Y_i + \tilde{Y}_i}{2},$$

then $E[Z_i] = E[Y]$. This means that Z_i is also an unbiased estimator, with lower variance.

Control variates: Suppose as usual that we would like to estimate the expected value $E[Y]$. On each simulation, we can generate one sample Y_i and another variable X_i , and thus draw a sequence of pairs (X_i, Y_i) . Assume that $E[X]$ is known, and define

$$Y_i(b) = Y_i - b(X_i - E[X]).$$

Algorithm 2 Antithetic variate Monte Carlo algorithm

```

1: procedure OPTION PRICE( $N, M$ )                                ▷ Path number  $N$ , time step  $M$ 
2:    $sum \leftarrow 0$ 
3:   for  $i = 1$  to  $N$  do
4:     generate  $W^1, W^2, W^3$  and calculate  $X, Y, Z$                 ▷ With (4.6)
5:     for  $j = 1$  to  $M$  do
6:       simulate paths of  $X_t^1, X_t^2$  with  $X, Y, Z$                 ▷ With (4.5a), (4.5b)
7:       simulate paths of  $\tilde{X}_t^1, \tilde{X}_t^2$  with  $-X, -Y, -Z$           ▷ With (4.5a), (4.5b)
8:       set paths of  $\xi_t^t$  and  $\tilde{\xi}_t^t$                                 ▷ With (3.8)
9:       simulate paths of  $S_t$  and  $\tilde{S}_t$                             ▷ With (4.5c)
10:    end for
11:     $sum \leftarrow sum + (F(S_T) + F(\tilde{S}_T)) / 2$                 ▷ Payoff function  $F(\cdot)$ 
12:  end for
13:  return  $sum/N$ 
14: end procedure

```

Note that $E[Y_i(b)] = E[Y_i]$, so

$$\frac{1}{M} \sum_{i=1}^M Y_i(b)$$

is an unbiased estimator of $E[Y]$. We can choose b to minimize variance of $Y_i(b)$:

$$\text{Var}[Y_i(b)] = \text{Var}[Y] - 2b\text{Cov}[X, Y] + b^2\text{Var}[X].$$

The optimal choice b^* is also the OLS coefficient regression of Y on X :

$$b^* = \frac{\text{Cov}[X, Y]}{\text{Var}[X]}.$$

And we will have

$$\text{Var}[Y_i(b)] = \text{Var}[Y](1 - \rho_{X,Y}^2).$$

It is desirable to choose a X that is strongly correlated with Y .

In practice, we would not be able to compute the value of b^* exactly and would estimate it in advance by simulation: choose n large and use

$$b^* = \frac{\sum_{i=1}^n (Y_i - \bar{Y})(X_i - E[X])}{\sum_{i=1}^n (X_i - E[X])^2}. \quad (4.7)$$

In other words, we would first run a pilot simulation (large n) to estimate b^* , and then use that fixed value throughout our main Monte Carlo simulation.

Algorithm 3 Control variate Monte Carlo algorithm

```

1: procedure OPTION PRICE( $N, M$ )                                ▷ Path number  $N$ , time step  $M$ 
2:    $sum \leftarrow 0$ 
3:   for  $i = 1$  to  $L$  do                                        ▷ Do pilot simulation first
4:     generate  $W^1, W^2, W^3$  and calculate  $X, Y, Z$               ▷ With (4.6)
5:     for  $j = 1$  to  $M$  do
6:       simulate paths of  $X_t^1, X_t^2$  with  $X, Y, Z$             ▷ With (4.5a), (4.5b)
7:       set paths of  $\xi_t^j$                                     ▷ With (3.8)
8:       simulate paths of  $S_t$                                   ▷ With (4.5c)
9:     end for
10:     $P \leftarrow F(S_t), Q \leftarrow H(S_t)$                 ▷ Payoff  $F(\cdot)$ , control variate  $H(\cdot)$ 
11:  end for
12:  Compute  $b^*$  for  $P, Q$                                        ▷ With (4.7)
13:  for  $i = 1$  to  $N$  do
14:    generate  $W^1, W^2, W^3$  and calculate  $X, Y, Z$               ▷ With (4.6)
15:    for  $j = 1$  to  $M$  do
16:      simulate paths of  $X_t^1, X_t^2$  with  $X, Y, Z$             ▷ With (4.5a), (4.5b)
17:      set paths of  $\xi_t^j$                                     ▷ With (3.8)
18:      simulate paths of  $S_t$                                   ▷ With (4.5c)
19:    end for
20:     $sum \leftarrow sum + F(S_T) + b^* (H(S_T) - E[H(S_T)])$ 
21:  end for
22:  return  $sum/N$ 
23: end procedure

```

For the Bergomi two-factor model, the simulation is path-dependent and it is difficult to choose good control variates. One can use the final spot price S_T as a control variate. This is a generic control variate, available to all products and models, but not always that efficient. The payoff function of a European call option is

$$F(S_T) = \max(S_T - K, 0),$$

resulting in the option price

$$C(S_0) = e^{-rT} E^Q[F(S_T)].$$

Further the underlying asset has known expected value $E^Q[S_T] = S_0 e^{rT}$, which yields the estima-

tion as:

$$C(S_0) = \frac{1}{N} \sum_{i=1}^N (\Phi(S_T^i) - b^*(S_T^i - S_0 e^{rT})).$$

Another better idea would be to use the sum of delta hedging P&L as the control variate. The delta hedge P&L control variate is formulated as follows:

$$\Delta_{CV} = \sum_{i=0}^{M-1} \frac{\partial C}{\partial S}(t_i, S_i, \hat{\sigma}) (S_{i+1} - S_i e^{r(t_{i+1}-t_i)}).$$

The delta is computed in the framework of Black-Scholes, with an arbitrary implied volatility $\hat{\sigma}$.

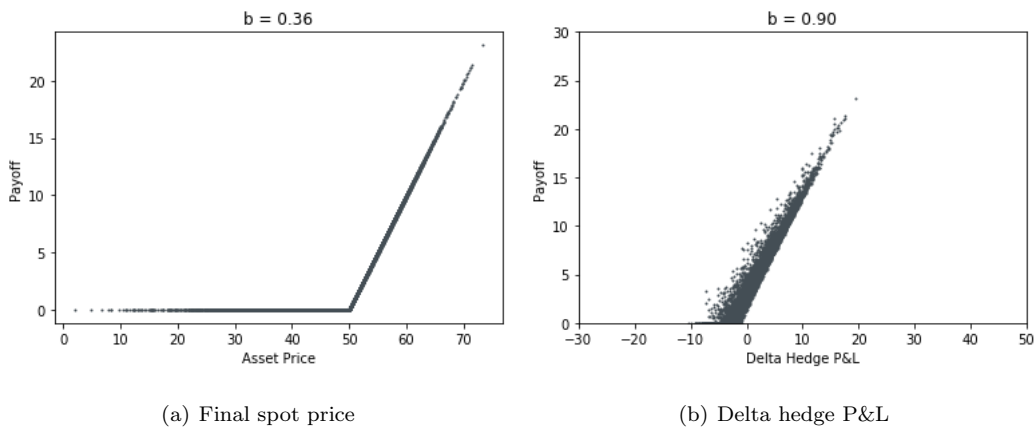


Figure 4: Correlation between control variate and option payoff for 1-year vanilla ATM call option with $S_0 = 50.0$, $r = 0.5\%$, in the two-factor model with parameters in Table 2, for different control variates. A flat VS volatility at 20% has been used.

We have compared the correlations between the two different control variates and the final payoff, as displayed in Fig.4. It indicates that the correlation between the delta hedge and the payoff is 0.9, and will lead to almost 5 times decreased standard error.

Gamma/theta P&L accrual: Bergomi [6] has proposed another technique based on a representation of the price of a European option in an arbitrary stochastic volatility model:

$$C(0, S_0) = C_{BS}(0, S_0, \hat{\sigma}) + \mathbb{E}^V \left[\int_0^T e^{-rt} \frac{S_t^2}{2} \frac{\partial^2 C_{BS}}{\partial S^2} (\sigma_t^2 - \hat{\sigma}^2) dt \right],$$

where \mathbb{E}^V denotes that the expectation is taken with respect to the dynamics generated by the stochastic volatility model at hand, whose instantaneous volatility process is σ_t . The equation expresses the price of a European option in an arbitrary stochastic volatility model as its price in the Black-Scholes model with implied volatility $\hat{\sigma}$ augmented by the discounted expectation of the integrated gamma/theta P&L evaluated with the Black-Scholes gamma - a natural representation from a trading point of view. We call $\hat{\sigma}$ is the risk-management volatility. We still need to simulate

the spot process, but S_t is only used to compute the Black-Scholes gamma. In our simulation, we use for $\hat{\sigma}$ the VS volatility for maturity T , and X_t^1, X_t^2, S_t are simulated at discrete times t_i : over the path, the second piece in the representation is evaluated as:

$$\sum_{i=0}^{N-1} e^{-rt_i} \frac{S_i^2}{2} \frac{\partial^2 C}{\partial S^2}(t_i, S_i, \hat{\sigma})(\xi_{t_i}^{t_i} - \hat{\sigma}^2)(t_{i+1} - t_i).$$

The pricing results and standard errors of Monte Carlo simulations with different techniques discussed above are illustrated in Table 6 and Fig.5, where we have used the parameters in Table 2. The VS volatilities are flat at 20% and the time step is set as 100.

Variance Reduction	Path Num	Price	Std Error	95% HW	Run Time
Original	250,000	3.59	0.83%	0.88%	10.06s
Antithetic Variate	125,000	3.60	0.53%	0.56%	9.02s
Control Variate (Underlying)	250,000	3.59	0.51%	0.54%	10.17s
Control Variate (Hedge)	250,000	3.59	0.18%	0.19%	15.07s
Gamma/Theta Accrual	250,000	3.60	0.18%	0.20%	13.59s

Table 6: Results of different variance reduction simulations

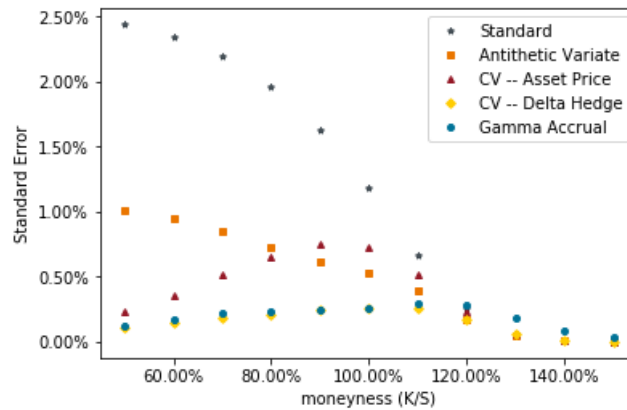


Figure 5: Standard errors of prices for 1-year vanilla ATM call option with $S_0 = 50.0$, $r = 0.5\%$, in the two-factor model with parameters in Table 2, for different Monte Carlo algorithms. A flat VS volatility at 20% has been used.

As expected, the simulation result with any variance reduction technique outperforms the one using standard Monte Carlo. However, the effectiveness is greatly reduced for the out-of-the-money options where it is barely more accurate than the standard Monte Carlo. It is also clear that the “control variate with delta hedge” and “gamma/theta P&L accrual” algorithms outperform the other techniques. However, even for the algorithm with the best performance, it still takes above

10 seconds to get a result within our tolerance of accuracy. It seems impossible for us to calibrate the Bergomi two-factor model with Monte Carlo simulations. We need to find alternatives.

In the following numerical simulations, without any further specification, we will utilize Monte Carlo with delta hedge as control variate to generate price and vanilla smile.

4.2.2 Flat VS term structure

For the case with correlated variance and underlying processes, the vanilla smiles of the Bergomi two-factor model with parameters in Table 2 have been illustrated in Fig.6. In the simulation, the term structure of VS volatilities is flat at 20%.

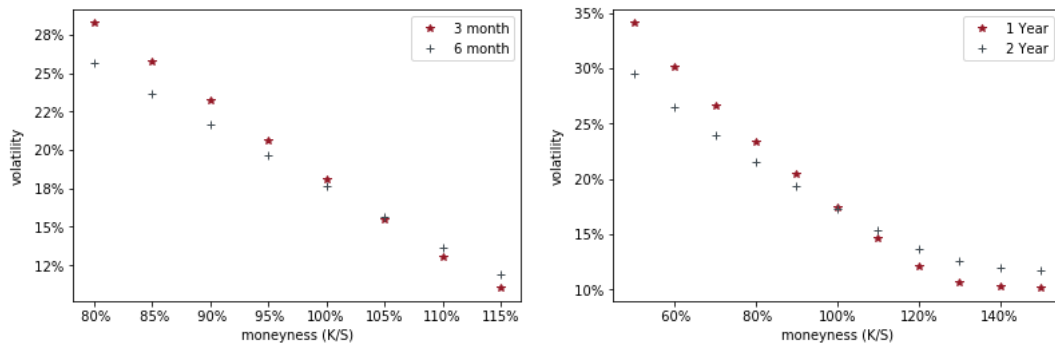


Figure 6: Smiles of vanilla ATM call option with $S_0 = 50.0$, $r = 0.5\%$, for the two factor model with parameters in Table 2. A flat VS volatility at 20% has been used.

The ATMF skew is displayed in Fig.7, as the difference of the implied volatilities for strikes $0.99F_T$ and $1.01F_T$, approximately $-0.02\mathcal{S}_T$, together with the order-one expression in (3.14). It is clear that the order-one expression is remarkably accurate to approximate the term structure of ATMF skew. We could utilize the order-one expression in our calibration.

And observe that in the Bergomi two-factor model, expression (3.14) for $\mathcal{S}_T^{\text{order } 1}$ does not involve the level of VS volatility: at order one in ϵ , the ATMF skew is unchanged if VS volatilities are rescaled by a common factor. Because of the accuracy of $\mathcal{S}_T^{\text{order } 1}$, we would expect this behaviour to persist in the simulated smile: this is illustrated in Fig.8. It shows the implied volatilities for different levels of VS volatility, for a one-year maturity vanilla ATM option. As is manifested, ATMF skew is practically independent on the level of VS volatility.

To finish the discussions in this part, we consider another case when the correlation of forward variance with underlying process vanishes: $\rho_1 = \rho_2 = 0$. It is well known that for uncorrelated spot and variances, the smile is symmetric in moneyness. The parameters used are listed in Table 7.

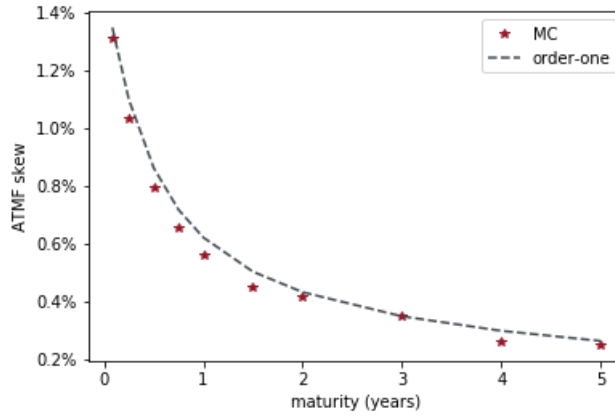


Figure 7: Term structure of simulated, as well as order-one approximated, ATMF skew of vanilla call option with $S_0 = 50.0$, $r = 0.5\%$, for the two-factor model with parameters in Table 2. A flat VS volatility at 20% has been used.

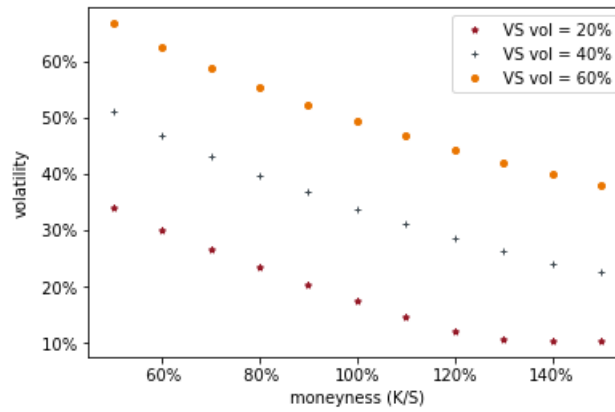


Figure 8: Smiles of vanilla ATM call option with $S_0 = 50.0$, $r = 0.5\%$, for the two factor model with parameters in Table 2. Different levels of flat VS volatility have been used.

ν	θ	κ_1	κ_2	ρ_{12}	ρ_1	ρ_2
174%	0.245	5.35	0.28	0.0%	0.0%	0.0%

Table 7: Values of parameters of the two-factor model II

Fig.9 displays the simulated smiles for four different maturities, for a flat term structure of VS volatilities at 20%. It is clear that the Bergomi two-factor model can capture the symmetric smile. However, the asymmetric skew is more common for equity index market.

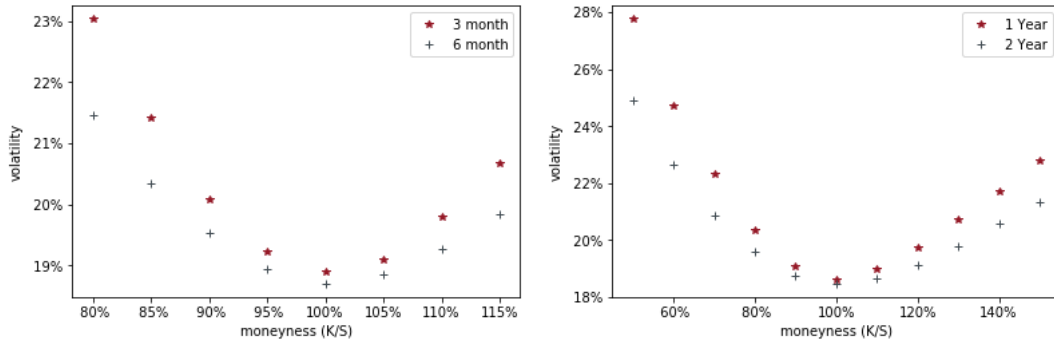


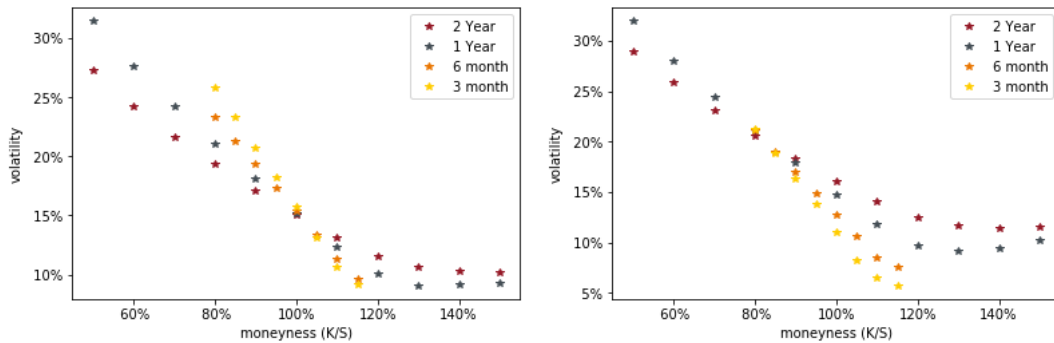
Figure 9: Smiles of vanilla ATM call option with $S_0 = 50.0$, $r = 0.5\%$, for the two factor model with parameters in Table 7. A flat VS volatility at 20% has been used.

4.2.3 Non-flat VS term structure

Normally, the assumption of the flat term structure for VS volatilities is not practical. In this section, we will use the following term structure:

$$\hat{\sigma}^{\text{VS}} = z_2 + (z_1 - z_2)e^{-z_3 T}, \quad (4.8)$$

where $z_1 = 12.03\%$, $z_2 = 21.80\%$, $z_3 = 2.166$, to simulate the vanilla smiles. The details of this parameterisation will be discussed in the following section. The parameters for the Bergomi two-factor model is the same as in Table 2 and the flat term structure is chosen as 17.8%, which is the VS volatility for one-year maturity. The simulation results are illustrated in Fig. 10 and Fig.11.



(a) Flat term structure for VS vol

(b) Non-flat term structure for VS vol

Figure 10: Smiles of vanilla ATM call option with $S_0 = 50.0$, $r = 0.5\%$, for the two-factor model with parameters in Table 2 – I

In Fig.10, the left graph shows vanilla smiles with flat term structure for VS volatilities. It is clear that the level of ATM volatility (100% moneyness) stays around the flat term structure value for all maturities, while for the non-flat term structure, from the right graph, the level of ATM

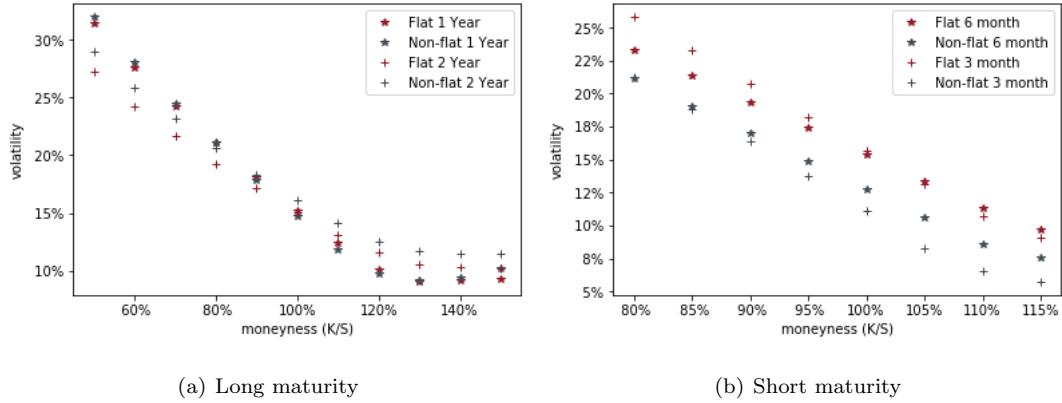


Figure 11: Smiles of vanilla ATM call option with $S_0 = 50.0$, $r = 0.5\%$, for the two-factor model with parameters in Table 2 – II

volatility (100% moneyness) varies according to the term structure value for different maturity, which can generate more flexible smiles in practice.

In Fig.11, we plot the smiles from flat term structure and non-flat term structure together and distinguish them by different maturities. It is noted that for short maturities, 3M and 6M, the effect of non-flat term structure is significant, adjusting the ATMF vol from around 17% to 12%. However, for longer maturities, there is only slight adjustment for ATMF volatility. It does make sense if we take into account the fact that the VS volatility/forward variance tend to stay stable for longer maturities. Moreover, it seems that there is no connection between the term structure and the ATMF skew, by observing the slope of the smile.

4.3 Approximation of the smile

In this part, we will conduct some numerical simulations to test the accuracy of the order-2 expansion of the implied volatility in (3.12). The implied volatilities from the order-2 expansion of the Bergomi two-factor model with parameters in Table 2 are displayed in Fig.12, together with the ones from Monte Carlo simulations. It is clear that the order-2 expansion works well mostly, especially for short maturity and around-the-moneyness options.

ν	θ	κ_1	κ_2	ρ_{12}	ρ_1	ρ_2
148%	0.231	6.00	0.25	0.0%	-70.0%	-35.7%

Table 8: Values of parameters of the two-factor model III

In Fig.13, the ATMF volatilities from order-2 expansion and Monte Carlo simulations are displayed for both a flat term structure of VS volatility as 17.8% and a non-flat one as in (4.8), with parameters in Table 2 and a new set in Table 8. The new set of parameters is chosen with a

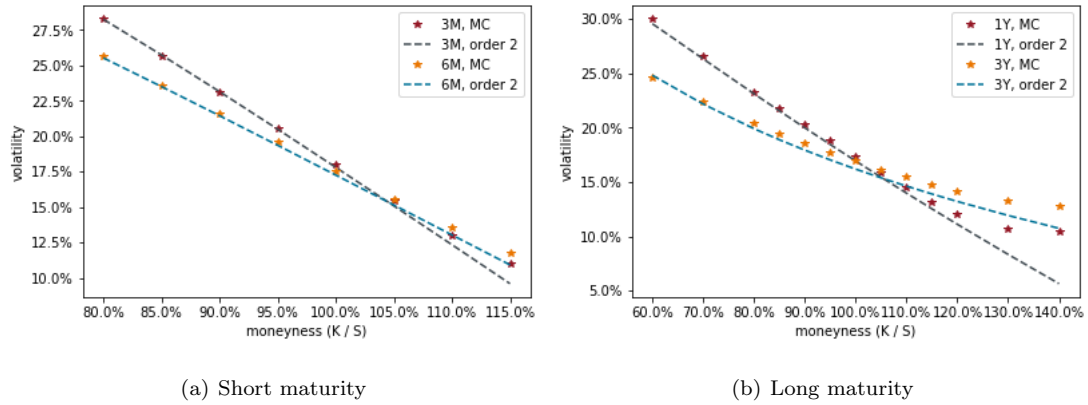


Figure 12: Smiles of vanilla call option with $S_0 = 50.0$, $r = 0.5\%$ from order-2 expansion and Monte Carlo simulation, for the Bergomi two-factor model with parameters in Table 2

lower instantaneous volatility of volatility.

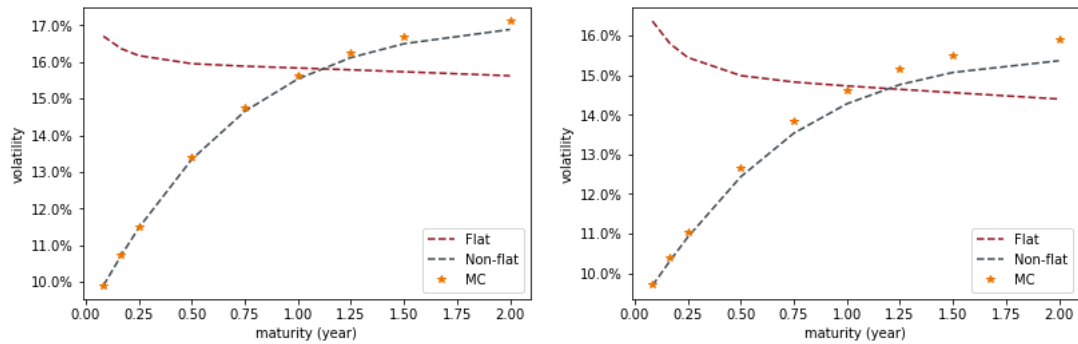


Figure 13: ATMF implied volatilities of vanilla call option with $S_0 = 50.0$, $r = 0.5\%$ from order-2 expansion and Monte Carlo simulations, with parameters in Table 2

The ATMF implied volatilities are extremely well estimated by the second order expansion. It is excellent for short maturities and small levels of volatility of volatility ν , but deteriorates at longer maturities. But for calibration, only the options around the money are utilized, where the second order expansion works well. Therefore, the approximation can be used in calibration, instead of Monte Carlo simulations. And as we have discussed, for a flat term structure of VS volatilities, it is clear that the ATMF implied volatilities stay around the flat value.

5 Practical Implementation

5.1 Sufficiency based on principle component analysis

In the previous section, we have introduced the general Bergomi N-factor model. For practical pricing purposes, we would like to drive the dynamics of the forward variance with a small number of factors. How many OU processes should we use and are enough to capture the desired dynamics? In this part, we will resort to Principle Component Analysis (PCA).

For a large dataset of correlated variables, principal components (PCs) enable us to summarize this dataset with a smaller number of variables that are able to explain most of the variability of the original dataset. The details can be found in any theoretical or applied statistical analysis textbook, see [23].

Here, one thing to be mentioned is that actually there are two equivalent ways conducting PCA. Simply put, the traditional PCA viewpoint requires one to compute the eigenvalues and eigenvectors of the covariance matrix, which is the product XX^T , where X is the centred data matrix of $n \times p$ size. Since the covariance matrix is symmetric, the matrix is diagonalizable, and the eigenvectors can be normalized such that they are orthonormal:

$$XX^T = WDW^T,$$

where W is a matrix of eigenvectors and D is a diagonal matrix with eigenvalues λ_i in the decreasing order on the diagonal. Moreover, $\lambda_i/(\lambda_1 + \dots + \lambda_p)$ is the proportion of the total variability of the data explained by the i th PC, and $(\lambda_1 + \dots + \lambda_k)/(\lambda_1 + \dots + \lambda_p)$ is the proportion of the variability explained by the first k PCs.

If we would like to retain k components, there is a trade-off between k large to explain more variability or k small to give a parsimonious representation. To assist in the choice of k , a diagram is often drawn, by plotting the points $(k, \lambda_k / \sum \lambda_i)$, or equivalently $(k, (\lambda_1 + \dots + \lambda_k) / \sum \lambda_i)$ and joining adjacent points by straight-line segments. Normally k is chosen such that

$$\lambda_1 + \dots + \lambda_k \gg \lambda_{k+1} + \dots + \lambda_p.$$

On the other hand, applying Singular Value Decomposition (SVD) to the data matrix X as follows:

$$X = U\Sigma V^T,$$

and attempting to construct the covariance matrix from this decomposition gives

$$\begin{aligned} XX^T &= (U\Sigma V^T)(U\Sigma V^T)^T \\ &= (U\Sigma V^T)(V\Sigma U^T). \end{aligned}$$

Since V is an orthogonal matrix,

$$XX^T = U\Sigma^2 U^T,$$

and the correspondence is easily seen.

Our first dataset includes over the counter quotes on VS rates on the S&P500 index extracted from Bloomberg. The data includes daily quotes with time to maturities 1, 2, 3, 6, 9, 12, 15, 18, and 24 months from November 4, 2008 to June 1, 2017, generating 2,133 observations for every maturity and a data matrix of $2,133 \times 9$ size. The centered data is denoted as X , i.e. column means have been subtracted from the original data and are equal to zero in X .

As demonstrated in Fig.14, PCA shows that the first principal component explains about 97% of the total variability of VS rates and can be treated as a level factor, while the second principal component explains an extra 2.33% and can be interpreted as a slope factor. This finding is not surprising, in that PCA of several other term structures, such as bond yields, produce similar results. But the first two factors are able to explain nearly all the variance of VS rates, i.e. 99%. Overall, PCA suggests that two factors are good enough to drive VS rates.

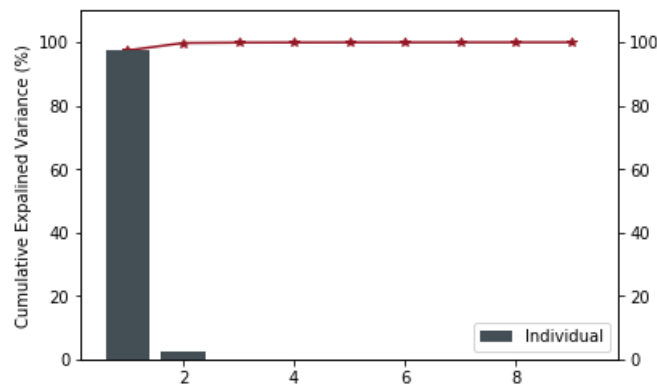


Figure 14: Explained variance by principle components

And while most term structures are upward sloping, they are sometimes U-shaped or downward sloping. We have reconstructed the variance swap curves with several principal components. Fig.15 shows that with either upward sloping or downward sloping, two principle components are enough to reconstruct the original curve, indicating that two OU processes are enough in Bergomi's forward variance model.

5.2 Calibration

The calibration of the Bergomi model is not based on the traditional standard: how it fits to the given market volatility surface. The calibration products include variance swaps and vanilla options. Several implementations of calibration will be described. As the term structure of the variance swap volatilities is important in constructing the volatility surface, we will first introduce two possible parameterisations.

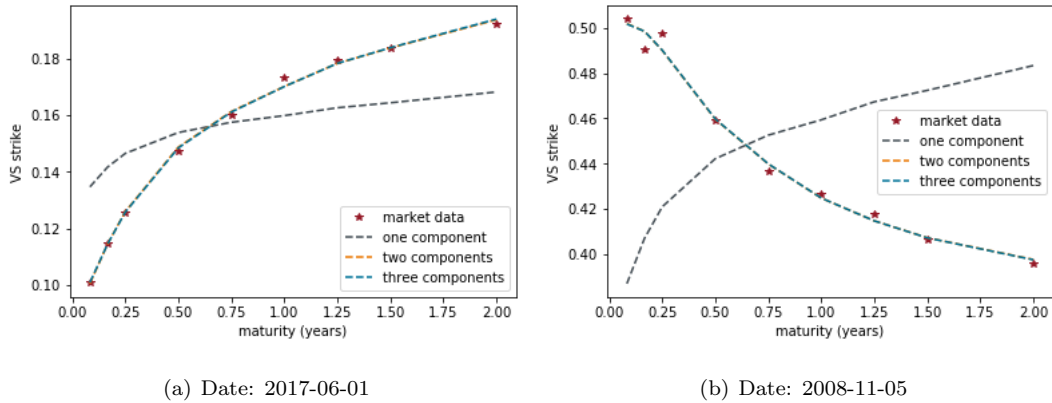


Figure 15: Variance swap curve reconstructed with principle components

5.2.1 Parameterisation of initial VS curve

In practise, the assumption of the flat term structure of forward variance/VS volatility does not hold, and the term structure is important in generating smiles. The initial values of forward variance curve can be calibrated on the market prices of variance swap contracts with:

$$\xi_0^t = \frac{d}{dt} ((\hat{\sigma}_0^t)^2 t), \quad (5.1)$$

where $\hat{\sigma}_0^t$ is the VS volatility for maturity t . Alternatively, the ξ_0^t can be chosen so as to recover the market prices of other instruments, such as ATMF volatilities of vanilla options for all maturities. In this section, we briefly discuss a few term-structure parameterisation schemes of variance swap curves.

From [11], a few schemes are given as: $\hat{\sigma}_0^t = z_1 + w(x)$ for some term-structure function w , and the following candidates have been considered:

$$\begin{aligned} w_1(x) &= z_2 \log(1 + t), \\ w_2(x) &= z_2 \sqrt{\epsilon + t}, \\ w_3(x) &= z_2 / \sqrt{\epsilon + t}, \\ w_4(x) &= z_2 \operatorname{atan}(\epsilon + t), \end{aligned}$$

where ϵ is a small value, which is set as $1e - 08$ in our simulation. However, all of these schemes only works well for either upward sloping or downward sloping. And even if we can vary our scheme according to the scenario, most of them will generate explosive forward variance curves. For example, with scheme w_1 , we have $\xi_0^t = \mathcal{O}(t)$, which is not desirable. In the long run, the values of initial variance swap curve and initial forward variance curve should approach to a constant level.

Therefore, it is reasonable to choose a parameterisation with asymptotic line. One choice is the

linearly mean-reverting parameterisation:

$$\hat{\sigma}_0^t = z_2 + (z_1 - z_2)e^{-z_3 t}, \quad (5.2)$$

for $z_1 \geq 0$ and $z_2, z_3 > 0$. z_1, z_2, z_3 stand for the short run variance, long run variance and the speed of mean-reversion respectively. The scheme is proposed as the functional for variance in [11], and it is a good choice to model volatility as well. Another candidate with asymptotic behaviour is the Gompertz function:

$$\hat{\sigma}_0^t = z_1 e^{-z_2 e^{-z_3 t}},$$

where z_1 is the asymptote, z_2 sets the displacement along the x-axis, i.e. time to maturity, and z_3 sets the growth rate.

From the discussions in the previous section, the term structure of the initial forward variance is very important in the volatility modelling with the Bergomi two-factor model. The initial forward variance curve ξ_0^t can be obtained analytically or via the numerical solution:

$$\xi_0^t = (\hat{\sigma}_0^t)^2 + t \frac{(\hat{\sigma}_0^{t+\epsilon})^2 - (\hat{\sigma}_0^{t-\epsilon})^2}{2\epsilon}$$

with $\epsilon = 1e - 08$.

The performances of fit for the two parameterisations are illustrated in Fig.16. Two different scenarios are considered: upward sloping and downward sloping initial VS volatility curves. It is clear that the two parameterisations work equally well for different term structure scenarios. Moreover, there are no significant differences between the extracted initial forward variance curves as well, and both of them present asymptotic behaviours, which are what we expect. Therefore, it is reasonable to fit the initial VS volatility curve with either of the parameterisations. With a more intuitive financial explanation, the linearly mean-reverting scheme is selected in the following implementations.

For the date 2016-04-18, which is also the date in calibration and pricing, the values of parameters used in mean-reverting scheme (5.2) are listed in Table 9. The short run variance is 12.03%, and the long run variance is 21.80%

z_1	z_2	z_3
12.03%	21.80%	2.166

Table 9: Values of parameters in linearly mean-reverting scheme on 2016-04-18

5.2.2 Market Data

The second dataset consists of market quotes on vanilla options on the S&P500 index from OptionMetrics.¹ The expiration date of index options is the Saturday following the third Friday of

¹ OptionMetrics, is a comprehensive database of historical market quotes, with best bid and best offer, for the entire US listed index and equity options markets.

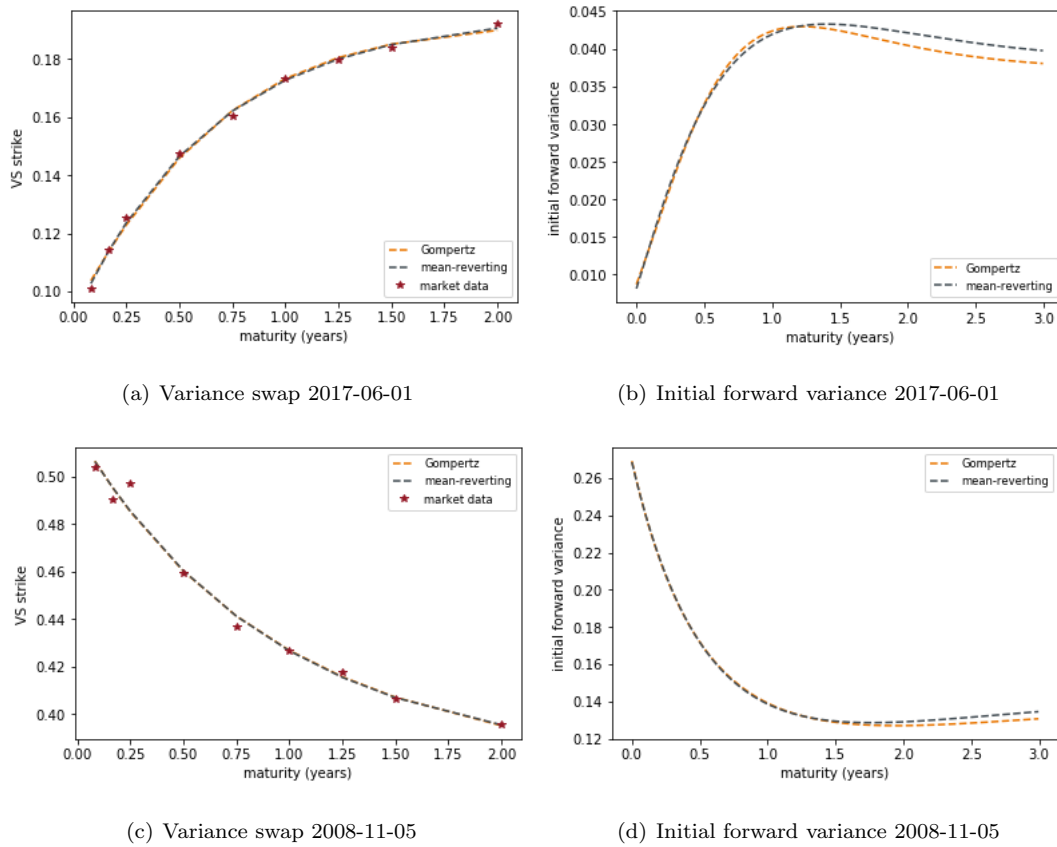


Figure 16: Performance of variance swap parameterisation schemes

the expiration month. And the index options generally may be exercised on last business day before the expiration date. Therefore, trading in index options will ordinarily cease on the business day (usually) preceding the expiration date. Part of market data of the options on S&P500 for 2016-04-18 is listed in Table 10.

strike	cp flag	date	ex date	best bid	best offer
300.0	C	20160418	20160520	1787.7	1790.8
300.0	P	20160418	20160520	0.0	0.25
400.0	C	20160418	20160520	1687.7	1690.9
400.0	P	20160418	20160520	0.0	0.25
...
2500.0	C	20160418	20160520	0.0	0.15
2500.0	P	20160418	20160520	408.0	411.2
300.0	C	20160418	20160617	1783.9	1787.1
300.0	P	20160418	20160617	0.0	0.05
...
3500.0	C	20160418	20181221	0.05	0.95
3500.0	P	20160418	20181221	1422.5	1432.2

Table 10: Market quotes of options on S&P 500 for 2016-04-18

From the discussions in the previous section, we could calibrate the correlations between forward variance and underlying processes to match the values of ATMF skew, measured as the difference of implied volatilities for strikes $0.99F_T$ and $1.01F_T$. Therefore, a proper interpolation scheme is necessary to construct the volatility surface. On the other hand, we also need zero rate / discount factor and spot price / forward price to calculate the price of exotics derivatives.

Calibrate discount factor and forward price Assume that the put-call parity holds for the European vanilla options:

$$C(K) - P(K) = DF \cdot (F - K),$$

where DF is the discount factor and F is the forward price. Therefore, we can utilize it to calibrate the discount factor and forward price for each maturity.

For the implementation, in practise, it is seldom to use all the instruments provided, considering that many of them have no open interest or trading volume. Instead, we can select a calibration window, say 30%, of all the instruments provided, and use them to calibrate our parameters. The idea is illustrated in Fig.17. The quoted prices in the “box” are used to conduct the calibration. For each maturity, we can obtain the respective discount factor and forward price through the following optimization problem:

$$\arg \min_{\Theta} \sum_K (C(K) - P(K) - DF \cdot (F - K))^2,$$

where $\Theta = \{DF, F\}$. A good choice for the initial guess of forward price used in the optimization would be the strike for which the call price and put price are equal. In Fig.17, it is illustrated as

the point of intersection of the call and put price curves.

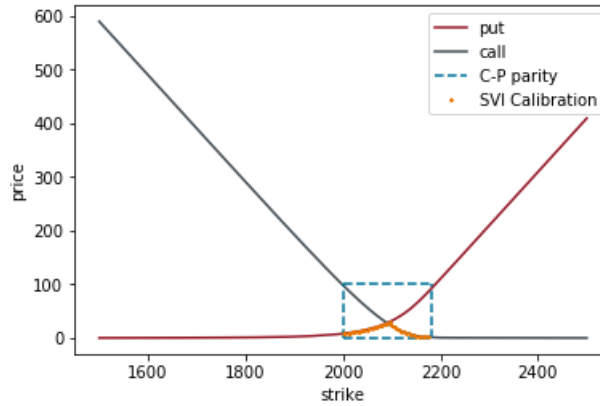


Figure 17: Illustration of the calibration window for expiry date 2016-05-20

The calibrated discount factor and forward price for different tenors on 2016-04-18 are listed in the following Table 11. These results will be used to simulate the Bergomi model and price other exotic instruments.

tenor (yr)	discount factor	forward price
0.088	0.999533	2090.27
0.164	0.999025	2087.41
0.241	0.998398	2085.65
0.337	0.997698	2081.99
0.414	0.997270	2079.24
0.663	0.994850	2072.21
0.759	0.994035	2071.12
0.912	0.992701	2065.88
1.162	0.990439	2059.74
1.660	0.985090	2048.70
2.677	0.969909	2030.96

Table 11: Values of discount factor and forward price on 2016-04-18

Construct volatility surface with raw SVI parameterisation Here, we consider an alternative, not based on some stochastic dynamics, but on a particular parameterisation, to construct the volatility surface with the provided market data. The SVI model was originally devised at Merrill Lynch in 1999 and subsequently publicly disseminated by Gatheral [12].

First, denote $C_{BS}(k, \sigma^2 t)$ as the Black-Scholes price of a European call option on S with strike $F_t e^k$, maturity t and volatility σ , where F_t is the forward price, and denote $\sigma_{BS}(k, t)$ as Black-

Scholes implied volatility. Then the total implied variance is $w(k, t) = \sigma_{\text{BS}}^2(k, t)t$. For a given parameter set $\Theta = \{a, b, \rho, m, \sigma\}$, the raw SVI parameterisation of total implied variance reads:

$$w(k; \Theta) = a + b \left\{ \rho(k - m) + \sqrt{(k - m)^2 + \sigma^2} \right\}, \quad (5.3)$$

where $a \in \mathbb{R}$, $b \geq 0$, $|\rho| \leq 1$, $m \in \mathbb{R}$, $\sigma > 0$, and the condition $a + b\sigma\sqrt{1 - \rho^2} \geq 0$, which ensures that $w(k, \Theta) \geq 0$ for all $k \in \mathbb{R}$.

The ease with which SVI can fit listed option prices led to its subsequent popularity with practitioners. In the implementation, we still select a calibration window, as illustrated by Fig.17. Due to the relatively high liquidity, only out of money calls and puts, with calibration window 15% for each, are utilized to calibrate the volatility surface. For each maturity, the optimization problem is:

$$\arg \min_{\Theta} \sum_k (w^{\text{SVI}}(k, \Theta) - w^{\text{MKT}}(k, t))^2.$$

The calibrated volatility surface is displayed in Fig.18. It is clear that the volatility surface generated by SVI parameterisation fits the market well. From the discussions in the previous section, ATMF skew and ATMF volatility are useful for the calibration of the Bergomi two-factor model, and generally, the SVI parameterisation tends to work well around-the-money.

Market volatility of volatility and ATMF skew To calibrate the forward variance process and to reflect the market's view in some way, the level and the term structure of the volatility of volatility, by the expression (3.10), should be consistent with historically observed ones.

As a measure of the volatility of the variance swap volatility with maturity τ for a given time scale Δ_t , Bergomi [4] defines

$$\text{Vol}(\hat{\sigma}_0^\tau) = \frac{1}{\sqrt{\Delta_t}} \text{Sd} \left[\log \left(\frac{\sqrt{V_{\Delta_t, \Delta_t + \tau}}}{\sqrt{V_0^{\Delta_t, \Delta_t + \tau}}} \right) \right],$$

where Sd is the standard deviation. For the calculation, we still use the data of over the counter quotes on variance swap rates on the S&P500 index extracted from Bloomberg as before. In this case, $\Delta_t = 1$ day. With the market quotes, the realized volatility of VS volatility is calculated as the sample standard deviation of the log-return of VS volatility:

$$\text{Vol}^2(\hat{\sigma}_0^\tau) = \frac{1}{n-2} \sum_{i=1}^{n-1} \left(\log(\hat{\sigma}_{0, i+1}^\tau / \hat{\sigma}_{0, i}^\tau) - \frac{1}{n-1} \sum_{i=1}^{n-1} \log(\hat{\sigma}_{0, i+1}^\tau / \hat{\sigma}_{0, i}^\tau) \right)^2 \times 252,$$

where n is the number of observations in our variance swap dataset.

After we construct the volatility surface from SVI parameterisation, the ATMF skew, measured as the difference of implied volatilities for the strikes $0.99F_T$ and $1.01F_T$, can be calculated directly through (5.3):

$$\sqrt{\frac{w(0.99; \Theta^*)}{T}} - \sqrt{\frac{w(1.01; \Theta^*)}{T}},$$

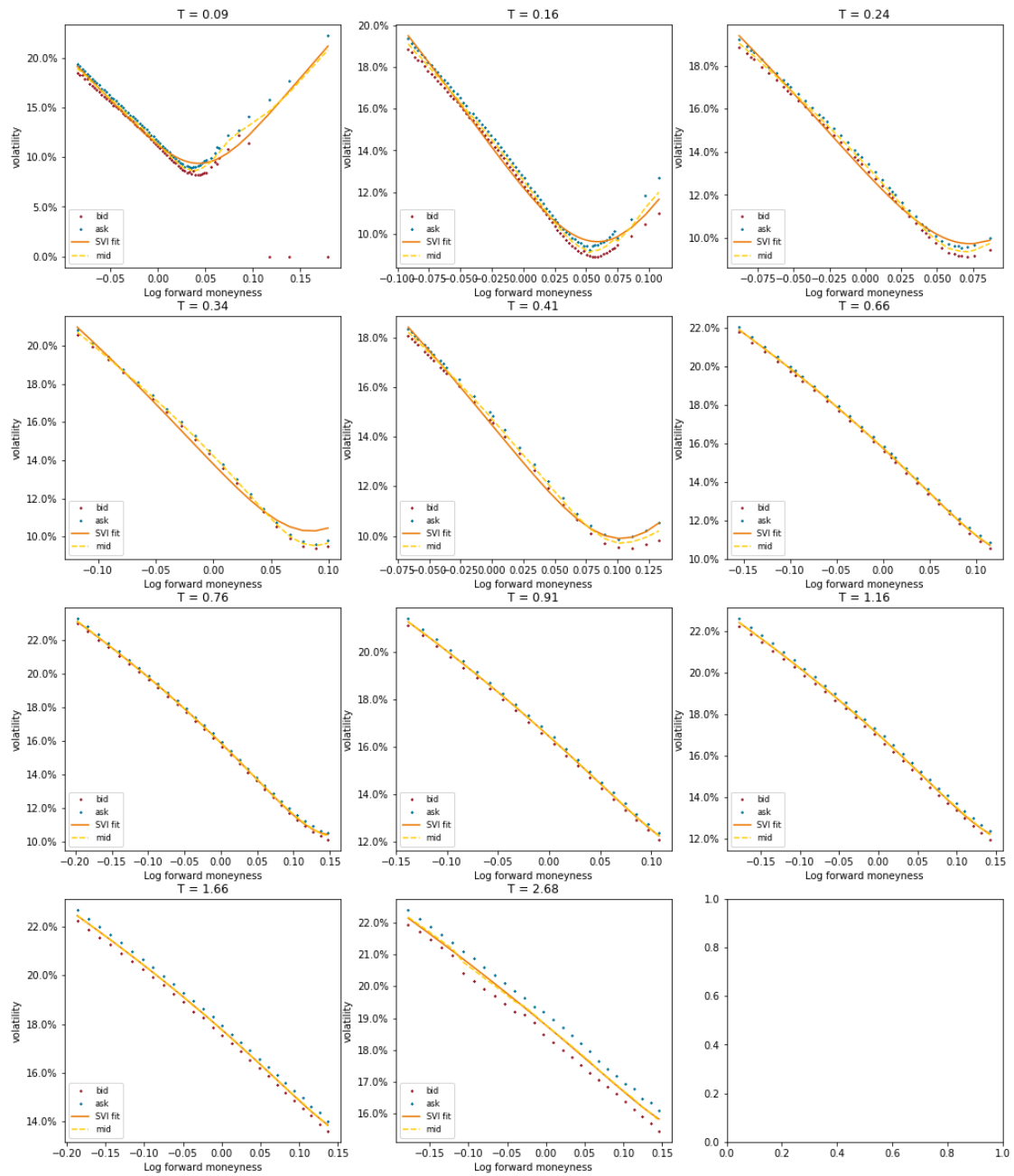


Figure 18: Market volatility surface with SVI parameterisation

where $\Theta^* = \{a^*, b^*, \rho^*, m^*, \sigma^*\}$ is the optimal parameters of SVI fit for maturity T . The market volatility of VS volatility and ATMF skew are displayed in Fig.19, together with the fitted benchmark functions (4.3).

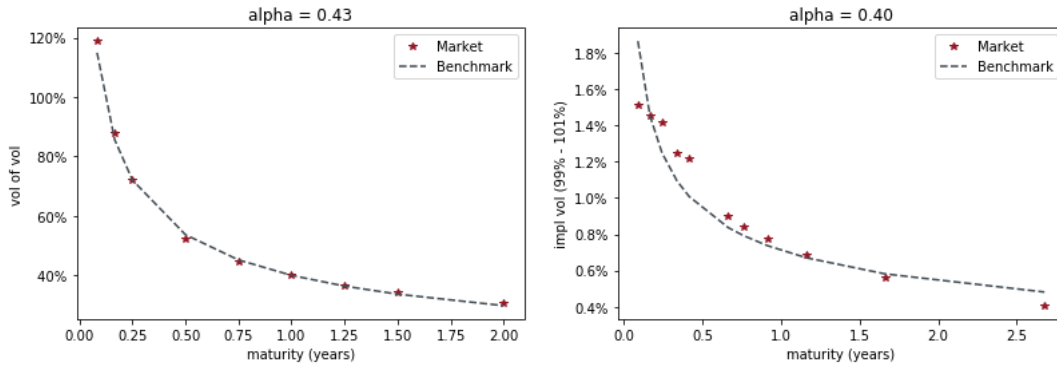


Figure 19: Market volatility of VS volatility and ATMF skew on 2016-04-18

Remark 5.1. From the right graph of Fig.19, the power-law benchmark only produces a general fit to the term structure of ATMF skew, which will have effect on the performance of calibration. The reason is that the S&P500 skew has been magnified during recent years, as is manifested in Fig.20. One potential explanation, see [30], is that dealer hedging of popular retirement products will affect the index greeks. We have mentioned in Section 1 that the US retirement market is dominated by both variable annuities and fixed indexed annuities. And the flows from dealers looking to hedge these positions is pushing up skew on the SPX. They usually buy put options and risk reversals to flatten their position, putting upward pressure on skew.

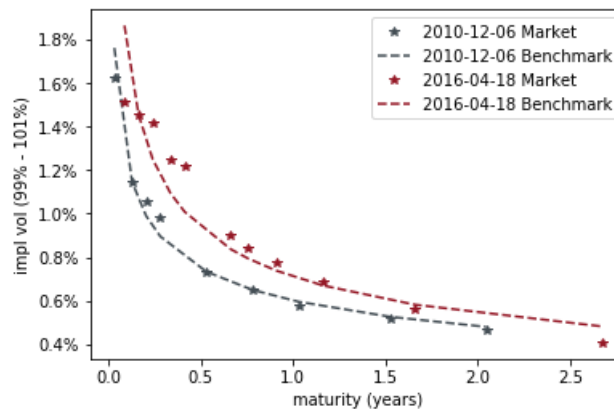


Figure 20: ATMF skew of S&P500 for date 2010-12-06 and date 2016-04-18

5.2.3 Calibration of Bergomi two-factor model

Now we are ready to implement our calibration for the Bergomi two-factor model. We will discuss three different ways of calibrating the model to the market. The non-flat term structure of VS volatilities in (5.2) has been used. And all the calibration procedures are tested for the date 2016-04-18. Another calibration example for the date 2010-12-06 is shown in Appendix B.

Separate calibration (Bergomi I) Bergomi [4] has introduced a two-step procedure for the calibration:

- Set dynamics for the forward variance process, i.e. choose the values for the parameters $\Theta = \{\nu, \theta, \kappa_1, \kappa_2, \rho_{12}\}$, by the following optimization problem:

$$\arg \min_{\Theta} \sum_T (\nu_0^T(\Theta) - \text{Vol}(\hat{\sigma}_0^T))^2,$$

where $\nu_0^T(\Theta)$ is calculated by (3.10) and $\text{Vol}(\hat{\sigma}_0^T)$ is the market volatility of VS volatility for maturity T .

- Calibrate the term skew, i.e. set the correlations $\Theta = \{\rho_1, \rho_2\}$ between the forward variance process and the asset process, by the following optimization problem:

$$\arg \min_{\Theta} \sum_T (\mathcal{S}_T^{\text{order } 1} - \mathcal{S}_T^{\text{MKT}})^2,$$

where $\mathcal{S}_T^{\text{order } 1}$ is calculated by (4.1) and $\mathcal{S}_T^{\text{MKT}}$ is the market ATMF skew for maturity T .

The separate calibration is very easy and fast to implement. We can calibrate the parameters in the Bergomi two-factor model step by step. The results of calibration are listed in Table 12. The time scales of the OU processes $1/\kappa_1$, $1/\kappa_2$ are clearly separated and can generate a volatility of VS volatility term structure that cannot be captured in a one-factor model.

ν	θ	κ_1	κ_2	ρ_{12}	ρ_1	ρ_2
177%	0.260	13.21	0.56	62.08%	-90.43%	-89.61%

Table 12: Values of parameters in Bergomi two-factor model with separate calibration

Although we can fit the volatility of VS volatility very well, the flexibility of the term structure of ATMF skew is restricted when the parameters for the forward variance process are set. As discussed before, the triplet $\rho_{12}, \rho_1, \rho_2$ must constitute a valid correlation matrix with restriction (4.4). As S&P500 skew has been largely magnified, it may be not possible to get such large level of skew, if we would like to recover the term structure of volatility of VS volatility at the same time.

The volatility of VS volatility and ATMF skew from Bergomi two-factor model with parameters from separate calibration have been displayed in Fig.21, together with the market values. It is

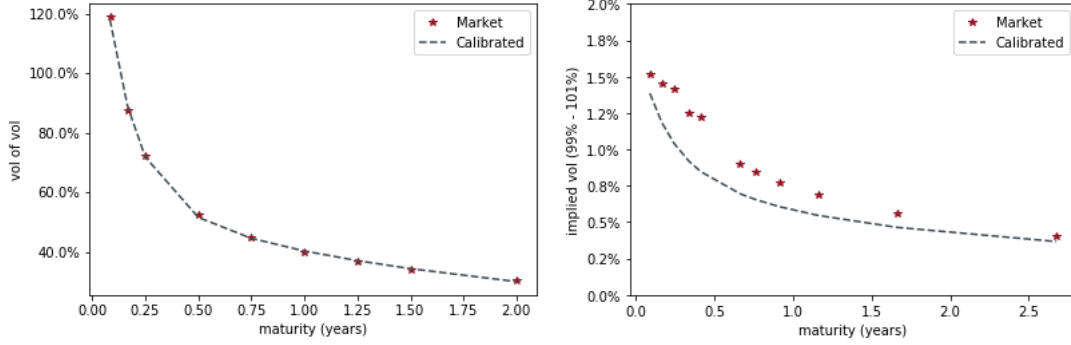


Figure 21: Volatility of VS volatility and ATMF skew with separate calibration

clear that we cannot match the ATMF skew very well, although the calibration works perfectly for volatility of VS volatility.

In Fig.22, the vanilla smiles are extracted using parameters from a separate calibration. The volatility surface produced by the model cannot reproduce the observed market, neither at the level of ATMF volatility nor at the ATMF skew.

Joint calibration (Bergomi II) From the previous discussions, the results from the separate calibration are not good enough to capture the market skew. Considering that the dynamics of skew is important in the pricing of cliquets, the joint calibration is introduced. In this implementation, all the parameters, $\Theta = \{\nu, \theta, \kappa_1, \kappa_2, \rho_{12}, \rho_1, \rho_2\}$, in the Bergomi two-factor model are calibrated simultaneously. As the ATMF skew is more important for pricing, a “weighted” error function is utilized:

$$\arg \min_{\Theta} \left(w_1 \sum_T \left(\frac{\nu_0^T(\Theta) - \text{Vol}(\hat{\sigma}_0^T)}{\text{Vol}(\hat{\sigma}_0^T)} \right)^2 + w_2 \sum_{\tilde{T}} \left(\frac{\mathcal{S}_{\tilde{T}}^{\text{order 1}} - \mathcal{S}_{\tilde{T}}^{\text{MKT}}}{\mathcal{S}_{\tilde{T}}^{\text{MKT}}} \right)^2 \right), \quad (5.4)$$

where $w_1 = 0.05, w_2 = 0.95$, with $\nu_0^T(\Theta)$ and $\mathcal{S}_{\tilde{T}}^{\text{order 1}}$ calculated by (3.10) and (4.1) respectively, and $\text{Vol}(\hat{\sigma}_0^T)$ as market volatility of VS volatility for maturity T and $\mathcal{S}_{\tilde{T}}^{\text{MKT}}$ as the market ATMF skew for maturity \tilde{T} . In the error function, relative errors are used to eliminate the issues of different scales for volatility of VS volatility and ATMF skew. In practical implementations, one can adjust the weight factor w_1 and w_2 .

The results of the joint calibration are listed in Table 13. Still, the time scales of the OU processes are clearly separated. The calculated volatility of VS volatility and ATMF skew from the model are displayed in Fig.23, together with the market values. It is noted that the ATMF skew matches the market well, but we have to sacrifice the fitness of volatility of VS volatility.

ν	θ	κ_1	κ_2	ρ_{12}	ρ_1	ρ_2
172%	0.118	4.69	0.01	75.36%	-98.93%	-84.12%

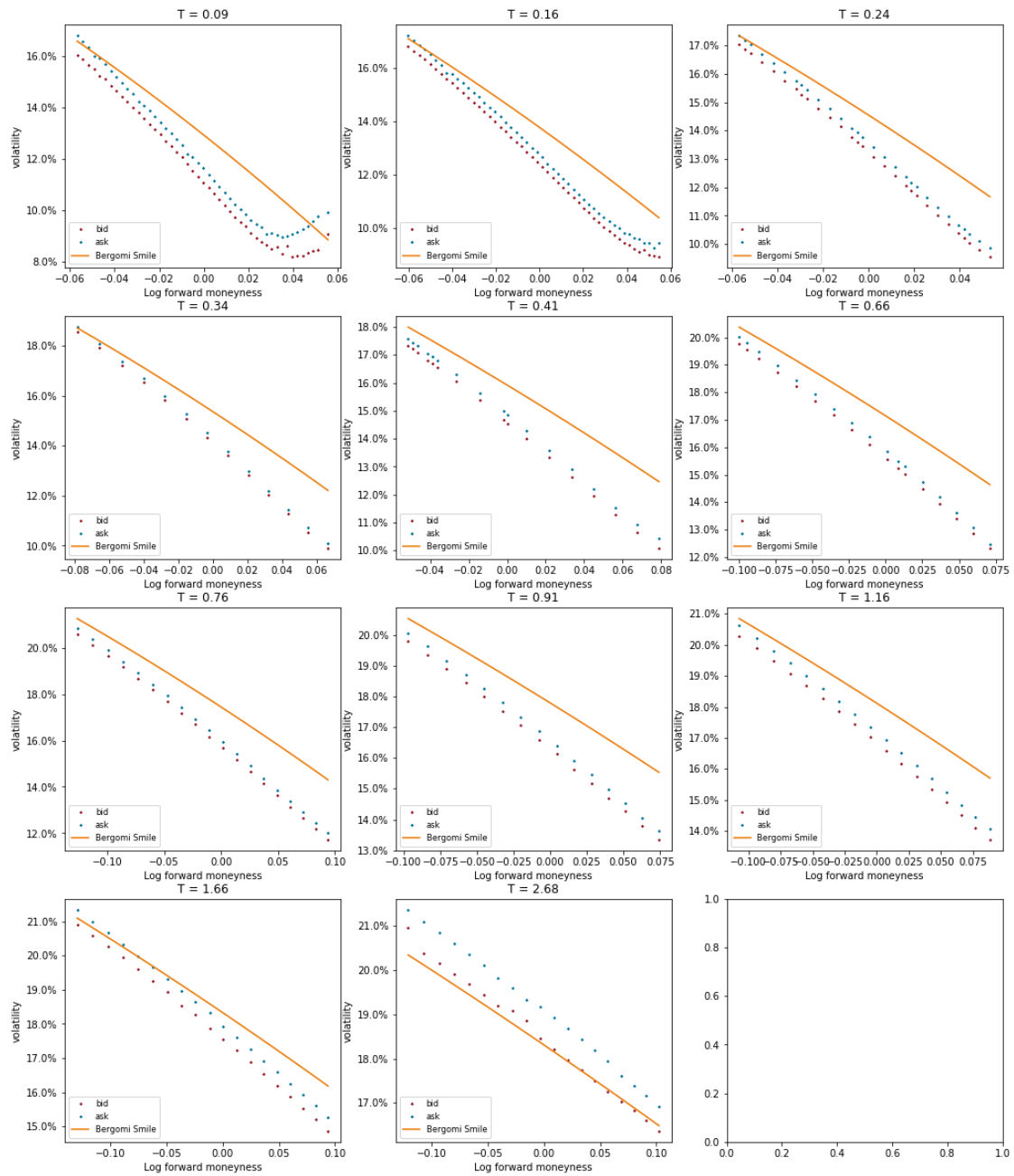


Figure 22: Vanilla smiles with separate calibration

Table 13: Values of parameters in Bergomi two-factor model with joint calibration

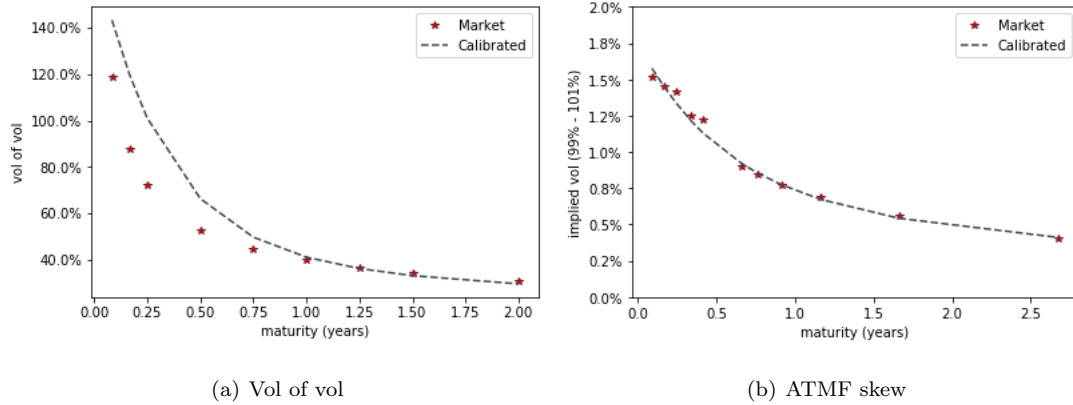


Figure 23: Volatility of VS volatility and ATMF skew with joint calibration

The smiles of implied volatilities with joint calibration are displayed in Fig.24. In most cases, the vanilla smile generated by Bergomi two-factor model with joint calibration can fit the market well. For short and long maturity, the ATMF skew is captured, but not the level of ATMF implied volatility. If the Bergomi model is assessed according to how it fits given market option prices across strikes and maturities, i.e., how it fits the volatility surface, then joint calibration produces more reasonable volatility surface than the separate calibration. And another flexibility in the joint calibration is that you can adjust the weights in (5.4), in order to get a customized calibration.

Calibrated to ATMF implied volatility (Bergomi III) Considering the drawbacks of joint calibration for the short and long maturity, another calibration implementation based on joint calibration will be introduced. From the previous discussions, the term structure of the initial VS volatility/forward variance has nothing or little effect on the ATMF skew. Therefore, we can calibrate a hypothetical term structure of VS volatility, according to (5.2) and the order-2 expansion of ATMF volatility (3.13a), to match the ATMF implied volatilities.

However, the initial VS volatility/forward variance curve is an input to the Bergomi two-factor model and the ATMF skew and ATMF volatility are jointly determined by the initial curve, with parameters z_1, z_2, z_3 and the other parameters $\nu, \theta, \kappa_1, \kappa_2, \rho_{12}, \rho_1, \rho_2$. It would not be a good idea to calibrate them, 10 in total, simultaneously. Alternatively, the method and results from the joint calibration can be utilized, and the new calibration is implemented through a two-step procedure:

- Calibrate $\nu, \theta, \kappa_1, \kappa_2, \rho_{12}, \rho_1, \rho_2$ simultaneously with the “weighted” error function (5.4) introduced in joint calibration.
- Fix the parameters from the joint calibration and calibrate the $\Theta = \{z_1, z_2, z_3\}$ in the initial

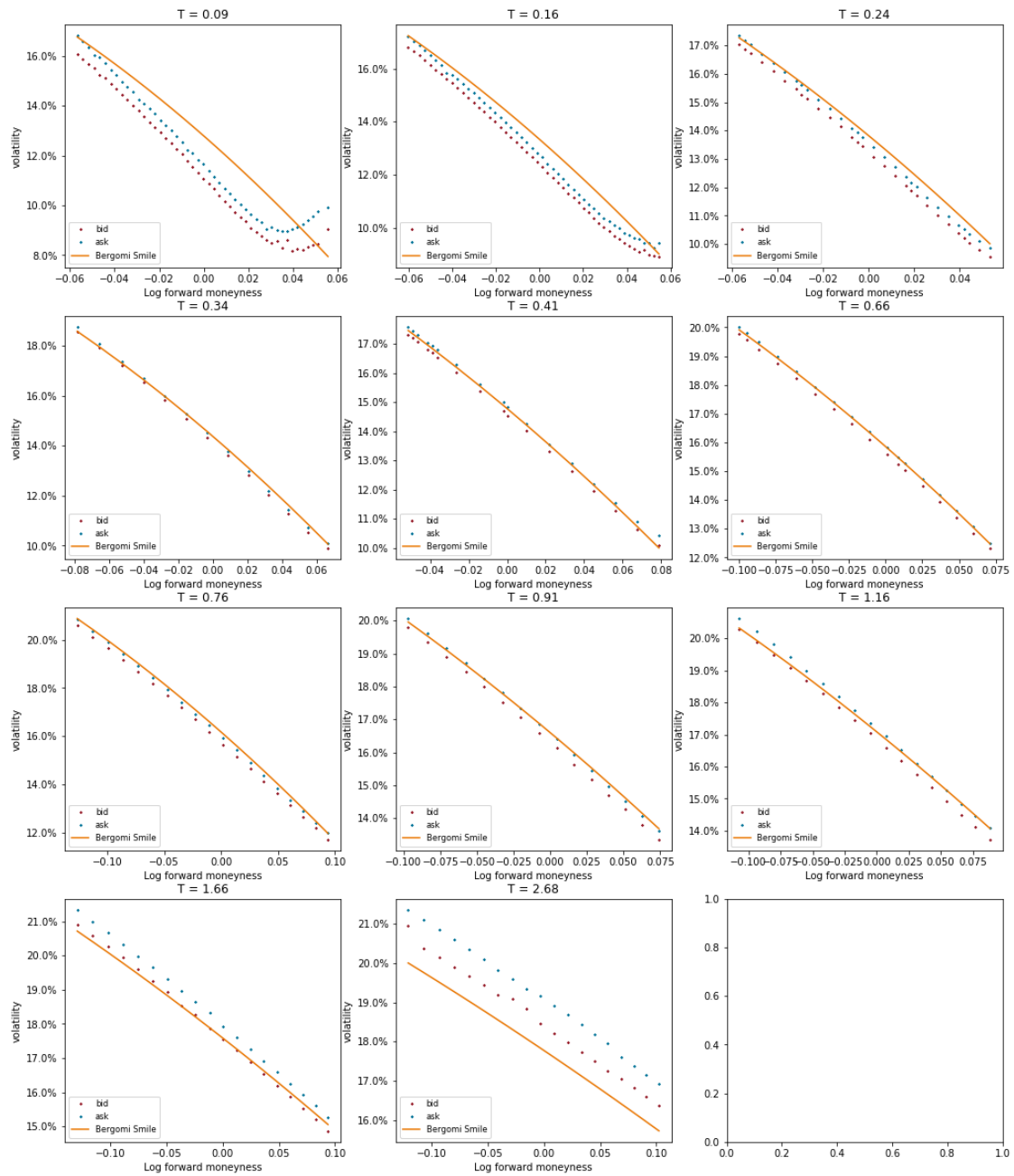


Figure 24: Vanilla smiles with joint calibration

VS volatility curve (5.2) by the following optimization problem:

$$\arg \min_{\Theta} \sum_T (\hat{\sigma}(F_T, T) - \hat{\sigma}^{\text{MKT}}(F_T, T))^2,$$

where $\hat{\sigma}(F_T, T)$ is calculated by (3.13a) and $\hat{\sigma}^{\text{MKT}}(F_T, T)$ is the market ATMF implied volatility for maturity T .

The new calibration is based on the results from joint calibration and takes more time than the previous two implementations. The results of calibration for the values of z_1, z_2, z_3 in the hypothetical term structure of VS volatility are listed in Table 14, and the values of other parameters for Bergomi two-factor model are the same as in Table 13. There are slight differences between the new z_1, z_2, z_3 and those used in previous two calibrations in Table 9. The calculated volatility of VS volatility and ATMF skew are displayed in Fig.25, together with the market values. Still, ATMF skew are fitted much better than the volatility of VS volatility, and the calibration can capture the observed levels of ATMF volatility.

z_1	z_2	z_3
10.40%	22.48%	1.94

Table 14: Values of parameters in VS volatility curve after calibrated to ATMF implied volatility

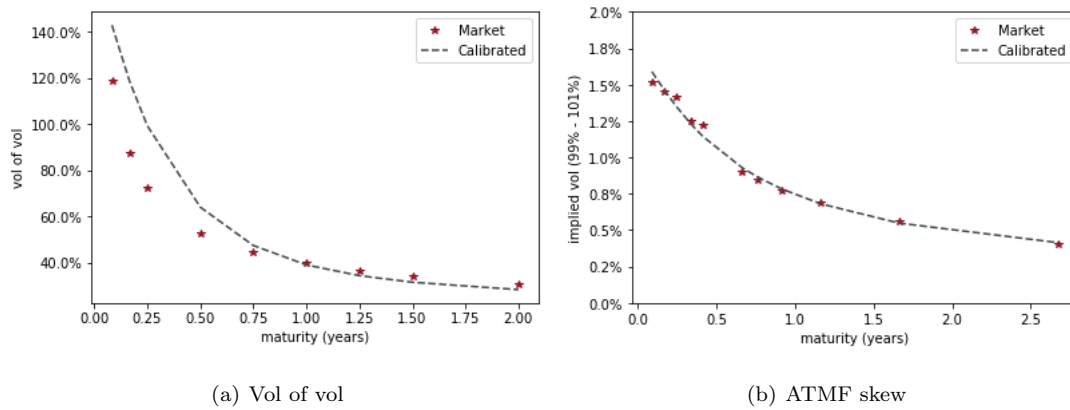


Figure 25: Volatility of VS volatility and ATMF skew after calibrated ATMF volatility

In Fig.26, the vanilla smiles after calibrated to the ATMF implied volatility are slightly better than the ones from the simple joint calibration, especially for short and long maturities. However, none of these three implementations can capture the curvature of smile well. It would be hard to obtain desirable curvature for every maturity if we calibrate the model only with ATMF implied volatility and ATMF skew. However, both joint calibration and the calibration to match ATMF volatility work well around-the-money.

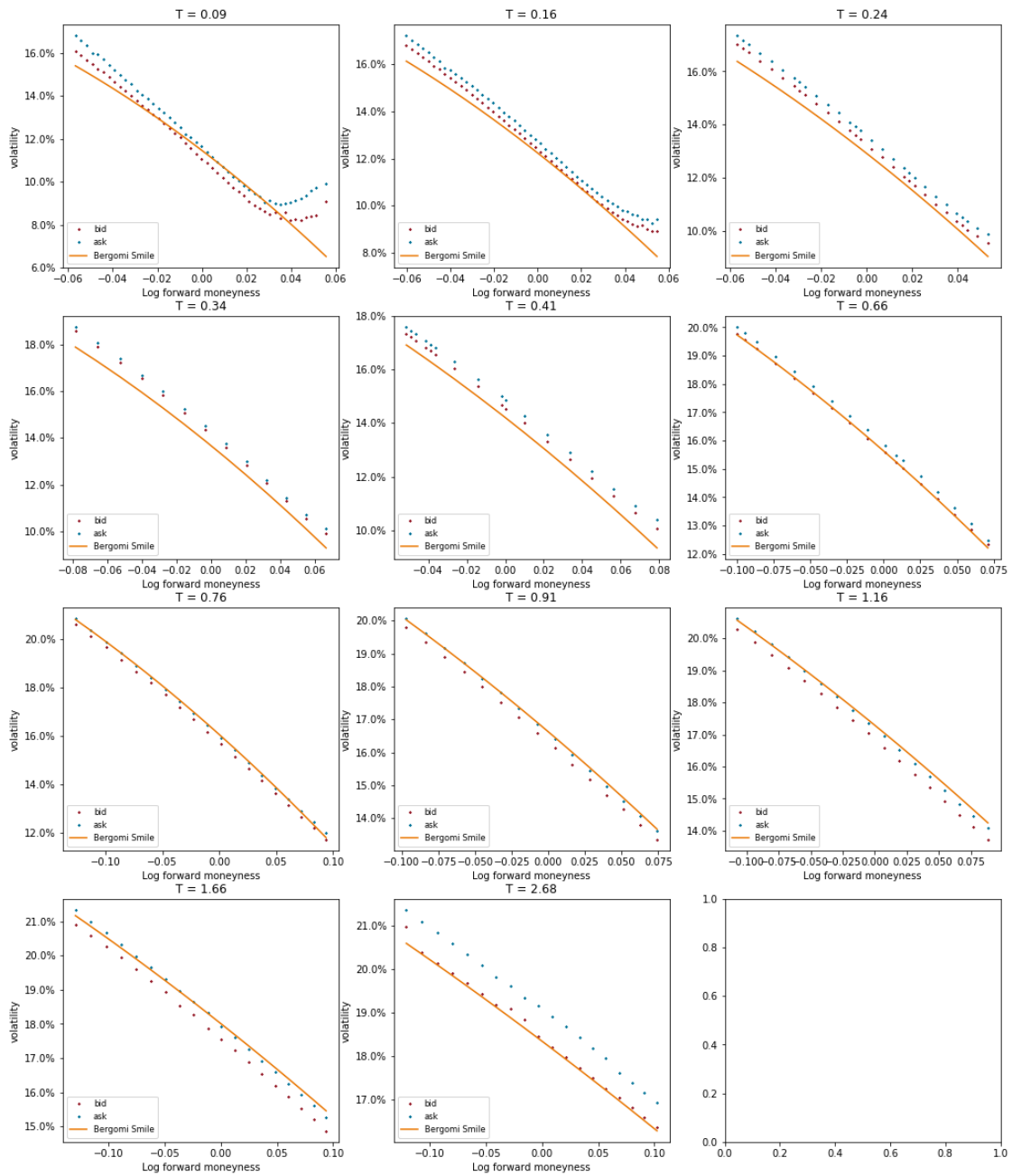


Figure 26: Vanilla smiles after calibrated to ATMF implied volatility

The summary of the three different calibrations is listed in Table 15. Calibrating a model amounts to deciding which instruments our exotic option price is a function of, along with the spot. If we calibrate our forward variance model to variance swaps, then these are the hedge instruments. Alternatively, if we calibrate so that the term structure of ATMF volatilities is recovered, we will use the corresponding vanilla options as hedges. Moreover, we should select the calibration method based on the payoff structure, i.e. more sensitive to the dynamics of volatility or skew.

	Vol of Vol	ATMF Skew	Vanilla Smile	Run Time
Bergomi I	✓	✗	✗	Shortest
Bergomi II	✗	✓	✓	Medium
Bergomi III	✗	✓	✓	Longest

Table 15: Comparison of three different implementations

5.2.4 Calibration of Heston model

The Heston model is assessed according to how it fits the given market option prices across all strikes and maturities. To calibrate the parameters $\Theta = \{\kappa, \theta, \nu, \rho, V_0\}$ in Heston model, the optimization problem is:

$$\arg \min_{\Theta} \sum_{T,K} (V^{\text{Heston}}(K, T) - V^{\text{MKT}}(K, T))^2,$$

where V^{Heston} is the price of the contract calculated from (2.5) and V^{MKT} is the market price. The calibration tries to fit the smiles for all maturities.

It is well known that the objective function of the Heston model calibration is not necessarily convex and may exhibit several local minimum, which complicates the estimation of the optimal parameters. The solution of the optimization might be dependent on the initial guess Θ_0 . Therefore, a good initial guess might be critical. From the previous discussions, we have already got a good sense of the market.

- Mean-reversion speed κ : the parameter should take positive value. However, it is not clear which upper value could be an appropriate bound. The initial guess is set as 5.
- Long-term variance θ : given the mean-reversion, the volatility of most financial asset rarely reaches levels beyond 100%. A good guess can be obtained from the (5.2) as z_2^2 , where z_2 is the long-run VS volatility.
- Volatility of variance ν : being a volatility, this parameter should exhibit positive values. However, the volatility of financial assets may change dramatically in short periods. The volatility itself is very volatile. We will use the ν in the Bergomi two-factor model as the initial guess for the Heston model.

- Correlation ρ : statistical correlation takes values from -1 to 1 . As previously discussed, the index market is highly skewed, indicating a strong negative correlation between the volatility and stock price. -80% is used as initial guess.
- Initial variance V_0 : similarly, the initial guess is set as z_1^2 from (5.2).

In the calibration, we try to fit the smiles for all maturities. Still, the calibration window is set as 30% , i.e. about $25 * 11 = 275$ options at hand. Even with good initial guess, the calibration for the Heston model is very slow. The results of the calibration are listed in Table 16.

κ	θ	ν	ρ	V_0
2.78	0.052	0.88	-0.85	0.015

Table 16: Values of parameters in the Heston model

Using these results, the model predicted vanilla smiles and their comparison with the market smiles are shown in Fig.27. The calibrated Heston model provides a general good match for the traded options. However, as the figure shows, it cannot fit the desired ATMF skew well.

5.3 Pricing

In this part, we use the calibrated model to price an accumulator, a reverse cliquet and a napoleon. We will compare the prices from different calibrated models and analyze the relative contribution of dynamics of volatility and skew. As an illustration, to highlight the importance of the dynamics of volatility and skew for the valuation of this type of options, we also consider the prices obtained from the Black-Scholes model. To this end, we use Monte Carlo simulations with 20 time steps for each month and 150,000 trials.

Table 17 and Table 18 display the prices corresponding to monthly cliquet options for maturity with 1 year and 3 years. Note that, as expected, the pricing errors associated with the Black-Scholes framework, with constant instantaneous volatility are extremely high for cliquet-style options. And the Heston model provides higher prices than Bergomi model, overestimating the related risks.

As discussed previously, cliquet-style options are very sensitive to the dynamics of skew. The fitness to ATMF skew can lead to important discrepancies in the pricing of exotic skew dependent derivatives. With the similar performance of the fitness to ATMF skew, Bergomi II and Bergomi III perform equally in pricing. The three methods of Bergomi calibration can provide an interval, regarded as the bid-ask spread, for the fair price of cliquet-style options.

	Bergomi I	Bergomi II	Bergomi III	Heston	Black-Scholes
Accumulator	2.31%	2.47%	2.47%	3.22%	1.25%
Reverse Cliquet	2.21%	2.87%	2.87%	3.78%	0.77%
Napoleon	1.45%	1.78%	1.78%	2.02%	1.03%

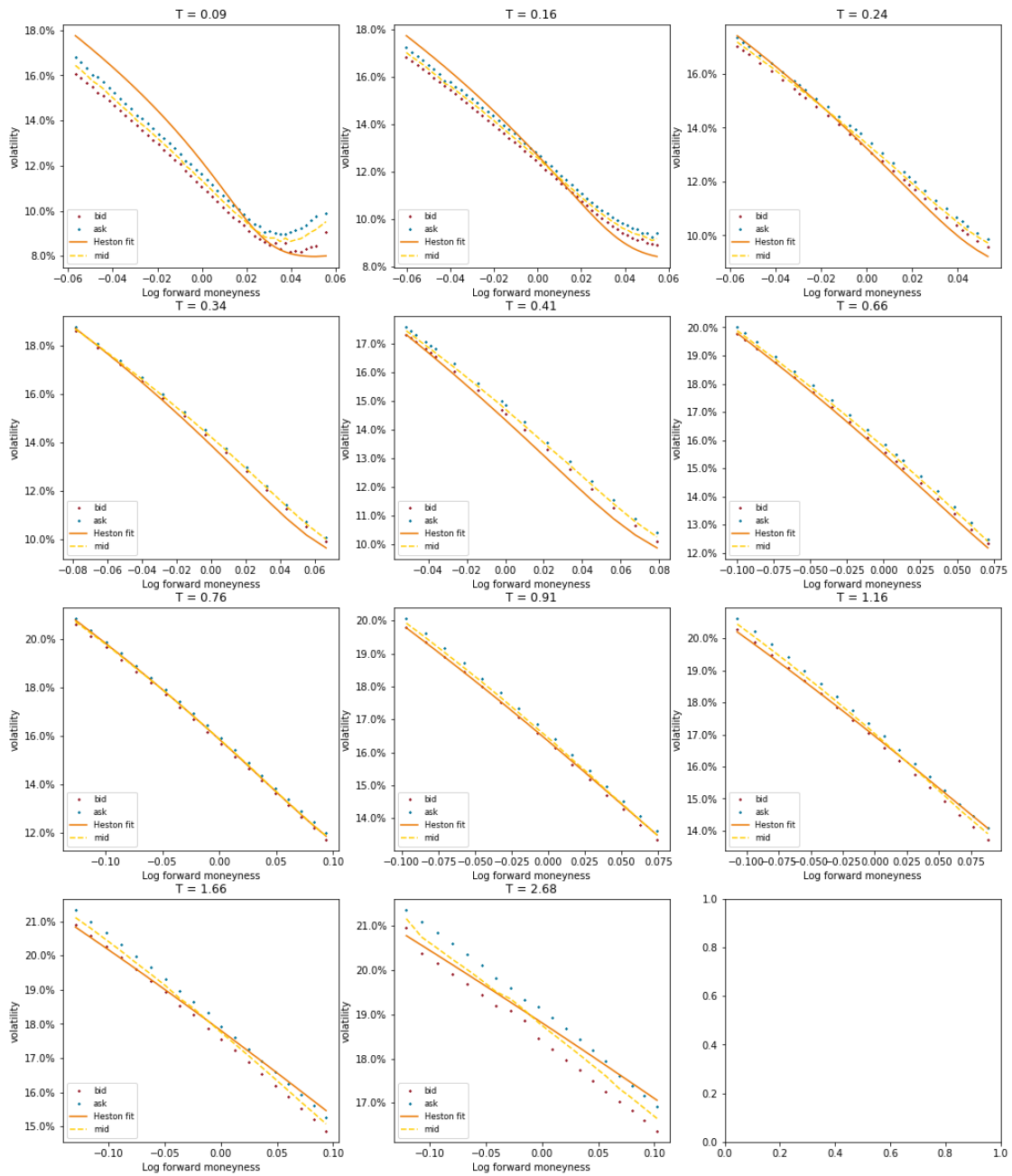


Figure 27: Vanilla smiles from the Heston model

Table 17: Model prices of cliquets with scenario I: B-S volatility 17.8%, maturity 1 year. Accumulator: monthly basis, floor 1%, cap 1%. Reverse cliquet: monthly basis, coupon 15%. Napoleon: monthly basis, yearly payout, coupon 8%

	Bergomi I	Bergomi II	Bergomi III	Heston	Black-Scholes
Accumulator	5.48%	6.00%	6.00%	8.12%	2.06%
Reverse Cliquet	3.22%	4.50%	4.24%	4.71%	0.27%
Napoleon	4.61%	5.48%	5.33%	5.67%	3.06%

Table 18: Model prices of cliquets with scenario II: B-S volatility 17.8%, maturity 3 years. Accumulator: monthly basis, floor 1%, cap 1%. Reverse cliquet: monthly basis, coupon 45%. Napoleon: monthly basis, yearly payout, coupon 8%

To test the impact of dynamics of skew on the pricing of cliquet-style options, another hypothetical pricing example has been conducted. In this case, we switch off the dynamics of skew by setting $\rho_1 = 0.0\%$ and $\rho_2 = 0.0\%$. The prices are listed in Table 19. It is noted that without correlations between the variance and underlying processes, the price of accumulator from the Bergomi model is just the Black-Scholes price. Almost all the contribution comes from the forward skew. Moreover, the volatility of volatility in the case when there is no skew has no material impact on the price of accumulator while it does when forward skew is switched on, by comparing the prices from three Bergomi implementations in Table 17 and 19. Again, for the Napoleon, the volatility of volatility accounts for most of the price, as differences between prices with and without forward skew are not very significant. As expected, the price of a reverse cliquet is sensitive to both volatility of volatility and forward skew.

	Bergomi I	Bergomi II	Bergomi III	Black-Scholes
Accumulator	1.24%	1.24%	1.25%	1.25%
Reverse Cliquet	1.11%	1.52%	1.56%	0.77%
Napoleon	1.26%	1.52%	1.54%	1.02%

Table 19: Model prices of cliquets with scenario III: B-S volatility 17.8%, maturity 1 year. Accumulator: monthly basis, floor 1%, cap 1%. Reverse cliquet: monthly basis, coupon 15%. Napoleon: monthly basis, yearly payout, coupon 8%

6 Conclusion And Further Work

In [29], Wilmott described Cliquet-style options as “the height of fashion in the world of equity derivatives”. As these contracts provide a downside protection while simultaneously offering a potential upside return, they, being insurance products, have become popular post-crisis. Cliquet-style options provide interesting opportunities for investors or arbitrageurs, especially in the current highly skewed market. Nevertheless, research literature on this topic is not widespread, neither in the theoretical modelling nor in the practical implementations. There is no agreement on the model for the pricing of these products.

To price them accurately, we should have an in-depth understanding of these contracts. From an investor’s perspective as well as the product issuer’s, the choice of the local cap and floor is crucial and changes the pricing results dramatically. And we have illustrated that these products are very sensitive to the dynamics of volatility, or dynamics of skew, or both. Due to these properties, great challenges will be faced by exotic traders when they try to price these path-dependent or forward-starting structures.

This thesis has mainly explored the Bergomi forward variance model. In order to capture the dynamics of volatility, it makes sense to model tradables such as variance swaps rather than non-tradables such as implied volatilities. And we confirmed that using principle component analysis of variance curves that two factors are sufficient. In contrast to the popular Heston model, it provides the flexibility to match the term structure of the volatility of volatility and the ATMF skew and allows the model user to directly control the behaviour of future smiles and hence properly price forward smile risk of cliquet-style exotic products.

The practical implementations have been discussed in a comprehensive way to test the calibration capabilities of these models in this thesis. We have abandoned the assumption of the flat term structure of variance swap curve used by Bergomi [4], [6] and validated the effectiveness of the linear mean-reverting parameterisation of the variance swap curve. The normal calibration of Bergomi two-factor model involves two levels of calibration. In the first step, we calibrate the parameters that generate the forward variance curve to match the term structure of the volatility of volatility. At the second step, we use the calibrated parameters, from the first step, and calibrate the correlation coefficients to control the term structure of the skew of the vanilla smiles.

However, as shown in the thesis, there are some constraints on the term structure of volatility of volatility and ATMF skew in the normal calibration of Bergomi two-factor model for the current highly skewed market. Therefore, we also considered other two implementations of the calibration. Nevertheless, these alternatives are not sufficient to perfectly recover the observed term structure of volatility of volatility and ATMF skew simultaneously, but still could improve the market fitting in comparison with a simple two-step calibration, especially for a highly skewed market. And it is noted that we should select the calibration method based on the payoff structure, i.e. more

sensitive to the dynamics of volatility or skew.

For the pricing of cliquet-style options, the dynamics of volatility and skew contribute a lot to the fair value. The pricing errors associated with the Black-Scholes framework are extremely high and the Heston model will overestimate the risk of the dynamics of volatility and skew. The Bergomi model can treat these risks accurately. And slightly different prices are obtained from the three calibration implementations and can be regarded as the bid-ask spread.

To improve the current work, more efficient calibration techniques should be developed. And one can bring in another factor and consider a three-factor model. Alternatively, as suggested by Bergomi [4] and [6], one can construct a constant elasticity of variance (CEV) model for the underlying, along with consistent log-normal two-factor dynamics for the forward variances term structure.

A Expansion of the implied volatility

Bergomi and Guyon [7] has derived an approximation of the smile produced by the forward variance model at second order in the volatility of volatility. They introduce a scaling factor ϵ for the volatilities of forward variances and derive that at second order in ϵ , the implied volatility for maturity T and strike K are exactly quadratic in log-moneyness:

$$\hat{\sigma}(K, T) = \hat{\sigma}(F_T, T) + \mathcal{S}_T \ln \left(\frac{K}{F_T} \right) + \frac{\mathcal{C}_T}{2} \ln^2 \left(\frac{K}{F_T} \right) + \mathcal{O}(\epsilon^3).$$

The ATMF volatility $\hat{\sigma}(F_T, T)$, the ATMF skew \mathcal{S}_T and curvature \mathcal{C}_T are given by:

$$\begin{aligned} \hat{\sigma}(F_T, T) &= \hat{\sigma}^T \left[1 + \frac{\epsilon}{4Q} C^{x\xi} + \frac{\epsilon^2}{32Q^3} (12(C^{x\xi})^2 - Q(Q+4)C^{\xi\xi} + 4Q(Q-4)C^\mu) \right], \\ \mathcal{S}_T &= \hat{\sigma}^T \left[\frac{\epsilon}{2Q^2} C^{x\xi} + \frac{\epsilon^2}{8Q^3} (4C^\mu Q - 3(C_{x\xi})^2) \right], \\ \mathcal{C}_T &= \hat{\sigma}^T \frac{\epsilon^2}{8Q^4} (4C^\mu Q + C^{\xi\xi} Q - 6(C^{x\xi})^2), \end{aligned}$$

where $Q = \int_0^T \xi_0^s ds$ and $\hat{\sigma}^T = \sqrt{\frac{Q}{T}}$, the VS volatility for maturity T , and $C^{x\xi}, C^{\xi\xi}, C^\mu$ summarize the joint spot/variance dynamics of the model at hand. $C^{x\xi}$ and $C^{\xi\xi}$ are integrals of the spot/variance and variance/variance covariance functions evaluated on the initial variance curve, and C^μ involves an extra degree of model-dependence as it depends on the derivative of $C^{x\xi}$ with respect to ξ . The details of $C^{x\xi}, C^{\xi\xi}, C^\mu$ are showed in this appendix.

First set:

$$\begin{aligned} \lambda_1(t, u, \xi) &= \alpha_\theta (2\nu) \xi^u \left((1-\theta) \rho_1 e^{-\kappa_1(u-t)} + \theta \rho_2 e^{-\kappa_2(u-t)} \right), \\ \lambda_2(t, u, \xi) &= \alpha_\theta (2\nu) \xi^u \left((1-\theta) \sqrt{1-\rho_1^2} e^{-\kappa_1(u-t)} + \theta \chi \sqrt{1-\rho_2^2} e^{-\kappa_2(u-t)} \right), \\ \lambda_3(t, u, \xi) &= \alpha_\theta (2\nu) \xi^u \theta \sqrt{(1-\chi^2)(1-\rho_2^2)} e^{-\kappa_2(u-t)}, \end{aligned}$$

with

$$\chi = \frac{\rho - \rho_1 \rho_2}{\sqrt{1-\rho_1^2} \sqrt{1-\rho_2^2}}.$$

Then λ_i all have the following common form:

$$\lambda_i(t, u, \xi) = \alpha_\theta (2\nu) \xi^u \left(w_{iX} e^{-\kappa_1(u-t)} + w_{iY} e^{-\kappa_2(u-t)} \right).$$

Secondly, by setting

$$\mathcal{I}(\alpha) = \frac{1-e^{-\alpha}}{\alpha}, \quad \mathcal{J}(\alpha) = \frac{\alpha-1+e^{-\alpha}}{\alpha^2}, \quad \mathcal{K}(\alpha) = \frac{1-e^{-\alpha}-\alpha e^{-\alpha}}{\alpha^2}, \quad \mathcal{H} = \frac{\mathcal{J}(\alpha)-\mathcal{K}(\alpha)}{\alpha},$$

we get

$$C^{x\xi} = \alpha_\theta (2\nu) \xi^{3/2} T^2 (w_{1X} \mathcal{J}(\kappa_1 T) + w_{1Y} \mathcal{J}(\kappa_2 T)),$$

$$C^{\xi\xi} = \alpha_{\theta}^2 (2\nu)^2 \xi^2 T^3 (w_0 + w_X \mathcal{I}(\kappa_1 T) + w_Y \mathcal{I}(\kappa_2 T) + w_{XX} \mathcal{I}(2\kappa_1 T) + w_{YY} \mathcal{I}(2\kappa_2 T) + w_{XY} \mathcal{I}((\kappa_1 + \kappa_2) T)),$$

with

$$\begin{aligned} w_0 &= \sum_{i=1}^3 \left(\frac{w_{iX}}{\kappa_1 T} + \frac{w_{iY}}{\kappa_2 T} \right)^2, \\ w_X &= -2 \sum_{i=1}^3 \frac{w_{iX}}{\kappa_1 T} \left(\frac{w_{iX}}{\kappa_1 T} + \frac{w_{iY}}{\kappa_2 T} \right), \\ w_Y &= -2 \sum_{i=1}^3 \frac{w_{iY}}{\kappa_2 T} \left(\frac{w_{iX}}{\kappa_1 T} + \frac{w_{iY}}{\kappa_2 T} \right), \end{aligned}$$

$$w_{XX} = \sum_{i=1}^3 \frac{w_{iX}^2}{\kappa_1^2 T^2}, \quad w_{YY} = \sum_{i=1}^3 \frac{w_{iY}^2}{\kappa_2^2 T^2}, \quad w_{XY} = 2 \sum_{i=1}^3 \frac{w_{iX} w_{iY}}{\kappa_1 \kappa_2 T^2},$$

and

$$C^{\mu} = \alpha_{\theta}^2 (2\nu)^2 \xi^2 T^3 (C_1^{\mu} + C_2^{\mu}),$$

with

$$C_1^{\mu} = \frac{1}{2} w_{1X}^2 \mathcal{H}(\kappa_1 T) + \frac{1}{2} w_{1Y}^2 \mathcal{H}(\kappa_2 T) - w_{1X} w_{1Y} \frac{\mathcal{J}(\kappa_2 T) - \mathcal{J}(\kappa_1 T)}{(\kappa_2 - \kappa_1) T},$$

$$C_2^{\mu} = w_X'' \mathcal{J}(\kappa_1 T) + w_Y'' \mathcal{J}(\kappa_2 T) + w_{XX}'' \mathcal{J}(2\kappa_1 T) + w_{YY}'' \mathcal{J}(2\kappa_2 T) + w_{XY}'' \mathcal{J}((\kappa_1 + \kappa_2) T),$$

and

$$w_X'' = \frac{w_{1X}^2}{\kappa_1 T} + \frac{w_{1X} w_{1Y}}{\kappa_2 T}, \quad w_Y'' = \frac{w_{1Y}^2}{\kappa_2 T} + \frac{w_{1X} w_{1Y}}{\kappa_1 T},$$

$$w_{XX}'' = -\frac{w_{1X}^2}{\kappa_1 T}, \quad w_{YY}'' = -\frac{w_{1Y}^2}{\kappa_2 T}, \quad w_{XY}'' = -\frac{w_{1X} w_{1Y}}{\kappa_1 T} - \frac{w_{1X} w_{1Y}}{\kappa_2 T}.$$

B Practical implementation example

In this part, to compare the results, we will demonstrate the practical implementations of calibration and pricing of cliquet-style options for another date, 2010-12-16.

The calibrated term structures of volatility of VS volatilities and ATMF skew for three different calibrations are displayed in Fig.28. It is noted that, in the less skewed market, all the three calibrations perform the same and can capture the dynamics of volatility and skew simultaneously. The vanilla smiles from calibrated Bergomi I, II, III and the Heston model are displayed in Fig. 29, Fig. 30, Fig. 31, Fig. 32 respectively. For the market at that time, there are no significant differences between the three calibrations of Bergomi model. Still, the Heston model cannot fit the market skew well for all maturities.

The pricing results of accumulator, reverse cliquet, napoleon, using Monte Carlo simulations, are listed in Table 20 and Table 21. It is noted that there are no significant differences in the pricing of accumulator as well.

	Bergomi I	Bergomi II	Bergomi III	Heston	Black-Scholes
Accumulator	2.21%	2.20%	2.21%	2.89%	1.21%
Reverse Cliquet	1.52%	1.47%	1.56%	2.41%	0.49%
Napoleon	0.97%	0.94%	1.00%	1.31%	0.62%

Table 20: Model prices of cliquets with scenario I: B-S volatility 20.8%, maturity 1 year. Accumulator: monthly basis, floor 1%, cap 1%. Reverse cliquet: monthly basis, coupon 15%. Napoleon: monthly basis, yearly payout, coupon 8%

	Bergomi I	Bergomi II	Bergomi III	Heston	Black-Scholes
Accumulator	5.06%	5.05%	5.06%	7.23%	1.91%
Reverse Cliquet	2.18%	2.29%	2.45%	1.83%	0.08%
Napoleon	3.47%	3.46%	3.60%	3.78%	1.83%

Table 21: Model prices of cliquets with scenario II: B-S volatility 20.8%, maturity 3 years. Accumulator: monthly basis, floor 1%, cap 1%. Reverse cliquet: monthly basis, coupon 45%. Napoleon: monthly basis, yearly payout, coupon 8%

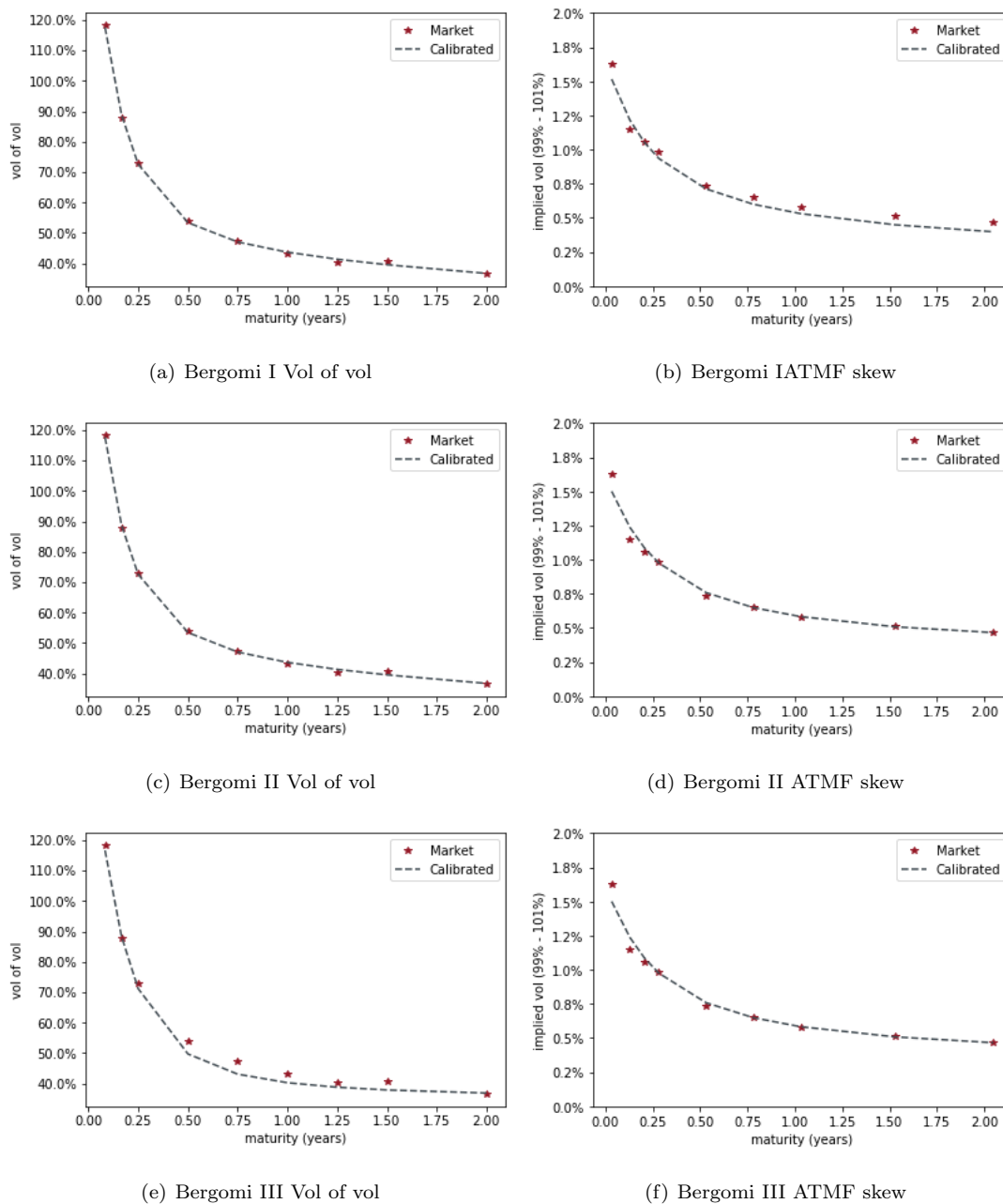


Figure 28: Volatility of VS volatility and ATMF skew with three different calibrations

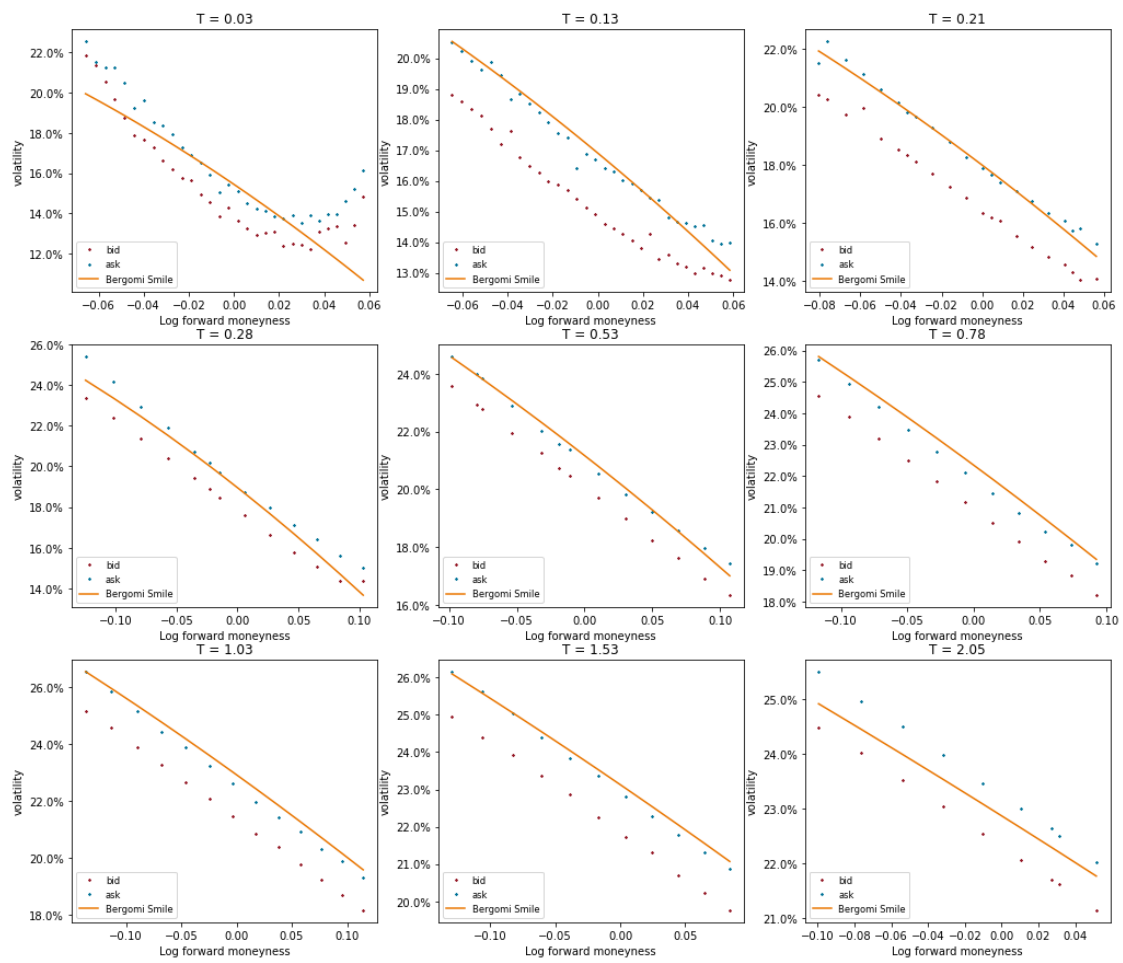


Figure 29: Vanilla smiles with separate calibration

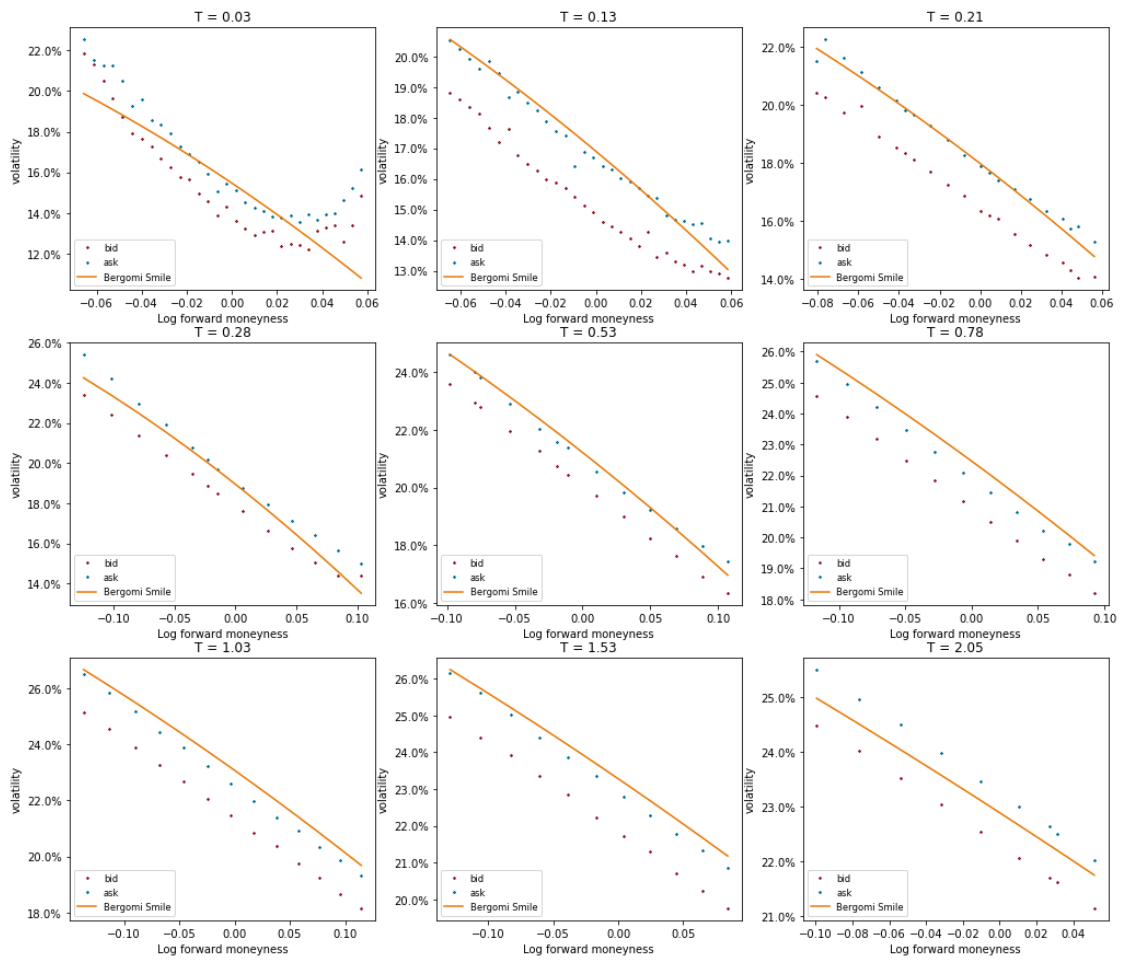


Figure 30: Vanilla smiles with joint calibration

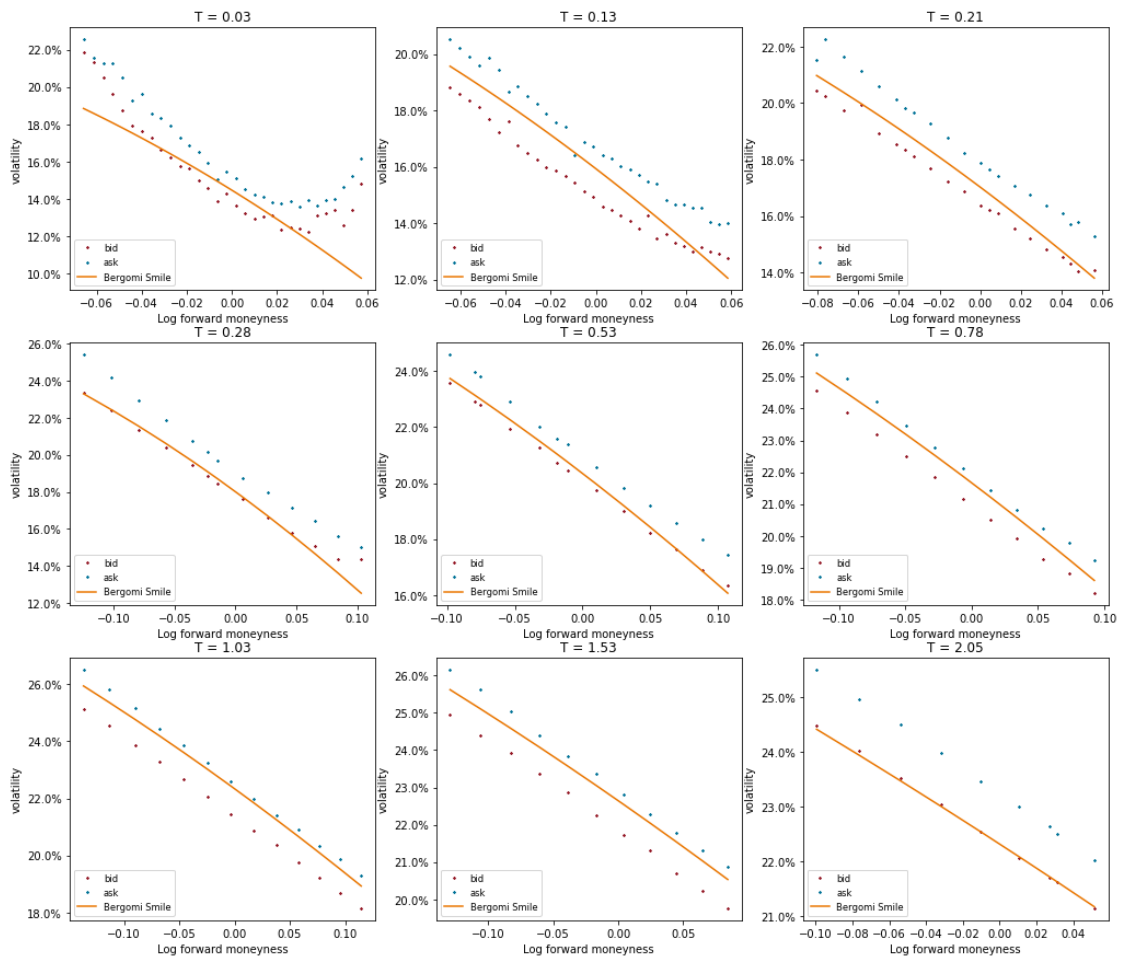


Figure 31: Vanilla smiles after calibrated to ATMF implied volatility

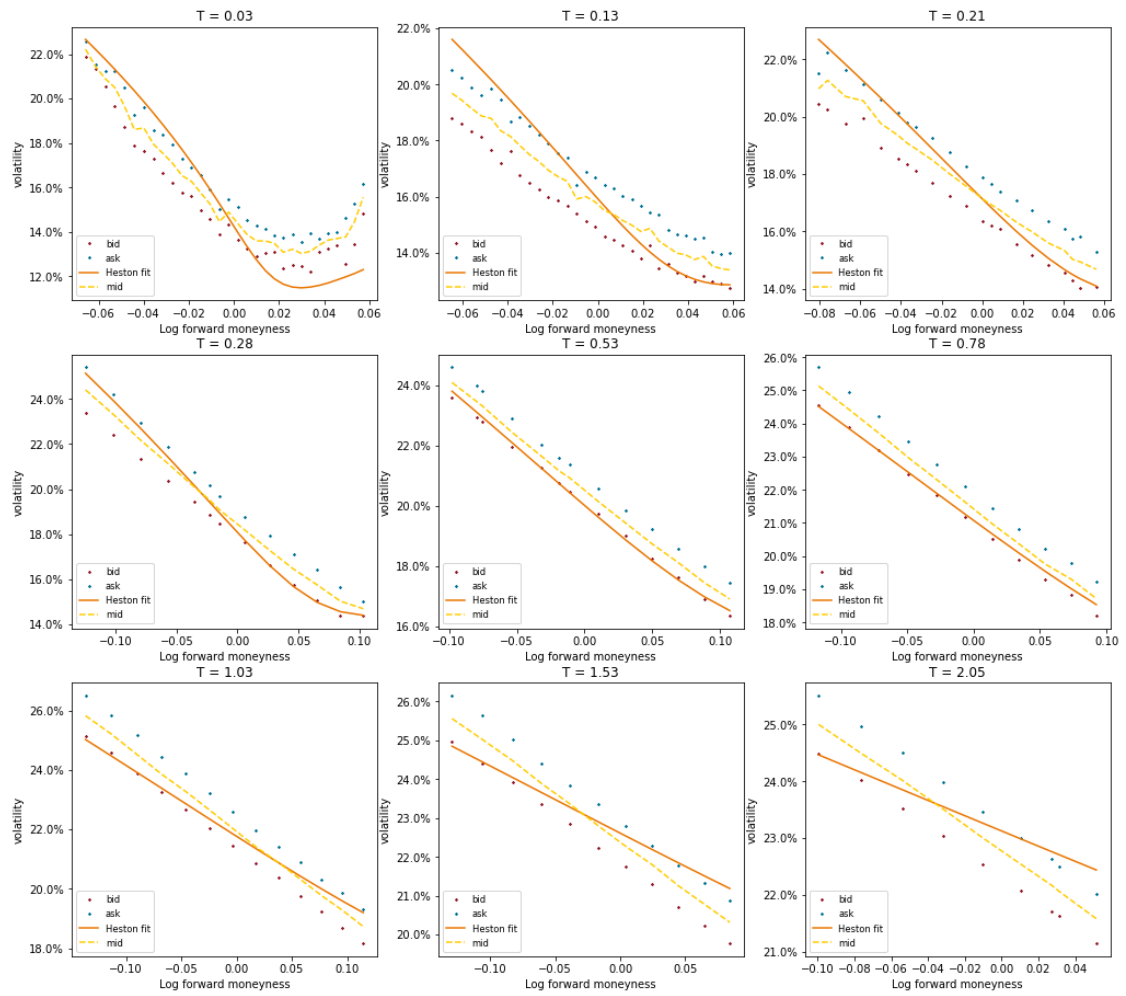


Figure 32: Vanilla smiles from the Heston model

References

- [1] Y. Ait-Sahalia, M. Karaman and L. Mancini. The term structure of variance swaps and risk premia. Available at SSRN: <https://ssrn.com/abstract=2136820>, 2015.
- [2] H. Albrecher, P. Mayer, W. Schoutens and J. Tistaert. The little Heston trap. *Wilmott Magazine*, 1, 83-92, 2006.
- [3] L. Bergomi. Smile dynamics. *Risk*, 9, 117-123, 2004.
- [4] L. Bergomi. Smile dynamics II. *Risk*, 10, 67-73, 2005.
- [5] L. Bergomi. Smile dynamics III. *Risk*, 10, 90-96, 2008.
- [6] L. Bergomi. *Stochastic Volatility Modeling*, CRC Press, 2015.
- [7] L. Bergomi and J. Guyon. The smile in stochastic volatility models. Available at SSRN: <https://ssrn.com/abstract=1967470>, 2011.
- [8] C. Bernard, W. V. Li. Pricing and hedging of cliquet contracts and locally-capped contracts. *SIAM Journal on Financial Mathematics*, 4(1): 353-371, 2013.
- [9] F. Black and M. Scholes. The Pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3): 637-659, 1973.
- [10] Gabriel G. Drimus. A forward started jump-diffusion model and pricing of cliquet style exotics. *Review of Derivatives Research*, 13(2), 2008.
- [11] H. Bühler. *Volatility Markets: Consistent Modelling, Hedging and Practical Implementation (Dissertation)*. Available at SSRN: <https://ssrn.com/abstract=1118245>, 2006.
- [12] J. Gatheral. A parsimonious arbitrage-free implied volatility parameterization with application to the valuation of volatility derivatives. *Presentation at Global Derivatives*, 2004.
- [13] J. Gatheral. *The Volatility Surface: A Practitioner's Guide*. John Wiley & Sons, 2006.
- [14] J. Gatheral. Developments in volatility derivatives pricing. *Presentation at Global Derivatives*, 2007.
- [15] J. Gatheral. Consistent modeling of SPX and VIX options. *Presentation at Bachelier Congress*, 2008
- [16] J. Gatheral and A. Jacquier. Arbitrage-free SVI volatility surfaces. *Quantitative Finance*, 14(1): 59-71, 2014.
- [17] J. Gatheral. Rough volatility: an overview. *Presentation at Advances in Financial Mathematics*, 2017.

-
- [18] J. Gatheral, T. Jaisson and M. Rosenbaum. Volatility is rough. Preprint available at arXiv:1410.3394.
- [19] P. Glasserman. *Monte Carlo Methods in Financial Engineering*. Springer, 2003.
- [20] J. Guyon. Stochastic volatility's orderly smiles. *Presentation at Fields Quantitative Finance Seminar*, 2013.
- [21] M. Haugh. Variance reduction methods I. Available at Monte Carlo Simulation: IEOR E4703, Columbia University, 2004.
- [22] S. L. Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies*, 6: 327-343, 1993.
- [23] G. James, D. Witten, T. Hastie and R. Tibshirani. *An Introduction to Statistical Learning*. Springer, 2013.
- [24] F. Kilin, M. Nalholm and U. Wystup. Numerical experiments on hedging cliquet options. *Journal of Risk*, 17(1): 85-103, 2014.
- [25] T. Lidebrandt. Variance reduction: three approaches to control variates. Available at <http://www.math.su.se/matstat>, 2007.
- [26] J. Marabel Romo. Pricing forward skew dependent derivatives. Multifactor versus single-factor stochastic volatility models. *Journal of Futures Markets*, 34: 124-144, 2014.
- [27] S. M. Ould Aly. Forward variance dynamics: Bergomi's model revisited. Preprint available at HAL Id: hal-00624812, 2011.
- [28] P. Warken. *Effective Pricing of Cliquet Options (Master's Thesis)*. Available at SlideShare: <https://www.slideshare.net>, 2015.
- [29] P. Wilmott. Cliquet options and volatility models *Wilmott Magazine*, 2: 78-83, 2002.
- [30] L. Woodall. Totally skewed: US annuity hedges magnify S&P volatility. Article available at Risk.net, 2016.