VOLATILITY MODELS AND OPTION PRICING

Dec 2022

Anass DAOU EL MAKANE, Hamza TARMADI

ÉCOLE POLYTECHNIQUE

IP PARIS

CONTENTS

1 OVERVIEW

Volatility is one of the major landmarks in financial mathematics, as an inherent measure of investment risk. As a standard deviation of the risky returns, it describes the overall performance of a stock market. Hence, higher volatility rates entail higher not only risks, but returns on average as well. [7]

Since the crash of October 1987, this concept has been of a paramount importance in order to explain the empirical behaviors, and sometimes anomalies in stock prices. Amongst these behaviors we consider:

- **Asymmetric volatility and leverage:** We can explain the negative correlation between volatility and stock prices with the leverage effect in equity markets. As the share price drops, debt percentage increase inducing thus higher volatility. From the investor point of view, as stock prices fall below a certain threshold, the only option for the investor is liquidate his assets entailing higher risks for the company.

- **Volatility smile:** The implied volatility is intuitively greater as the stock prices get far from the strike value (in or out the money). More explanations of this effect account for the financial sentiments from a behavioral point of view.

- **Volatility clustering:** Empirically, volatility rates exhibit strong autoregressive properties. This behavior is an important one, as it motivated new volatility models to replicate it.

There are some other advanced and market specific phenomena such as the mean reversion property, the power law of decay or even jump diffusion.
Throughout this article, we are going to explore the most widely used volatility models as well as their option pricing frameworks in order to capture as many empirical behaviors as possible.

One of the earliest attempts to provide a mathematical modeling of stock prices is due to **Bachelier** in 1900 [7]. Under $(\Omega, \mathcal{F}, \mathcal{P})$:

$$dS_t = \mu.dt + \sigma.dW_t^{\mathbf{P}}$$

This model couldn't satisfy the empirical observations, since a stock price is always non-negative, and more specifically, the returns and the prices should not be correlated.

Historically, the model that showed the most consistencies with what is shown in practice is the **Black-Scholes** model.

$$dS_t = S_t.(\mu(t, S_t).dt + \sigma(t, S_t).dW_t^{\mathbf{P}})$$

To begin with, we will study the constant and time dependent volatility models. Then, we will show their inconsistencies with empirical stock prices and derive local and stochastic approaches to study volatility.

2
BLACK-SCHOLES VOLATILITY:

As a first assumption, we assume that the volatility and drift are constant. Hence, the stochastic differential equation the becomes:

$$dS_t = S_t.\mu.dt + S_t.\sigma.dW_t^{\mathbf{P}}$$

A solution in this case would be the GBM with a constant risk free rate of return r:

$$\begin{cases} S_t = S_0.e^{(\mu - \frac{\sigma^2}{2}).t + \sigma.W_t^{\mathbf{P}}} \ (Risky\ asset) \\ B(t) = B_0.e^{r.t} \ (Risk\ free\ bonds) \end{cases}$$

From here, we apply the **Girsanov** theorem for a maturity T, with a deterministic change of measure $\lambda = \frac{\mu - r}{\sigma}$ such that $\frac{d\mathbf{Q}}{d\mathbf{P}} = e^{-\lambda.W_T^{\mathbf{P}} - \frac{1}{2}.\lambda^2.T}$ since λ is deterministic. Now under $(\Omega, \mathcal{F}, \mathbf{Q})$ where we denote $\tilde{S}_t$ the discounted stock price.

$$\frac{d\tilde{S}_t}{\tilde{S}_t} = (\mu - r).dt + \sigma.dW_t^{\mathbf{P}}$$
$$\implies \frac{d\tilde{S}_t}{\tilde{S}_t} = \sigma.dW_t^{\mathbf{Q}} \ (1)$$

Hence, the stock price is a **Q**- martingale with a log-normal distribution. From here, let $(X_t)_{0\leq t\leq T}$ be a self financing portfolio and $(\theta_t)_{0\leq t\leq T}$ an admissible portfolio strategy.

$$dX_t = \theta_t.\frac{dS_t}{S_t} + (X_t - \theta_t).r.dt$$

Applying Ito's Lemma to $\tilde{X}_t := X_t e^{-rt}$ the discounted wealth process we see that:

$$d\tilde{X}_t = e^{-rt}\theta_t.\frac{d\tilde{S}_t}{\tilde{S}_t}$$

The discounted wealth of is **Q**-a local martingale, and a super martingale for an admissible portfolio.

The pricing formula can be shown using the no arbitrage valuation:

$$\mathcal{C}(T,t,K,S_t) = e^{-r.(T-t)}.\mathbf{E}^{\mathbf{Q}}[(S_T - K)^+|\mathcal{F}_t]$$

Overall, there are two common methods to extract informations about the volatility from empirical data.

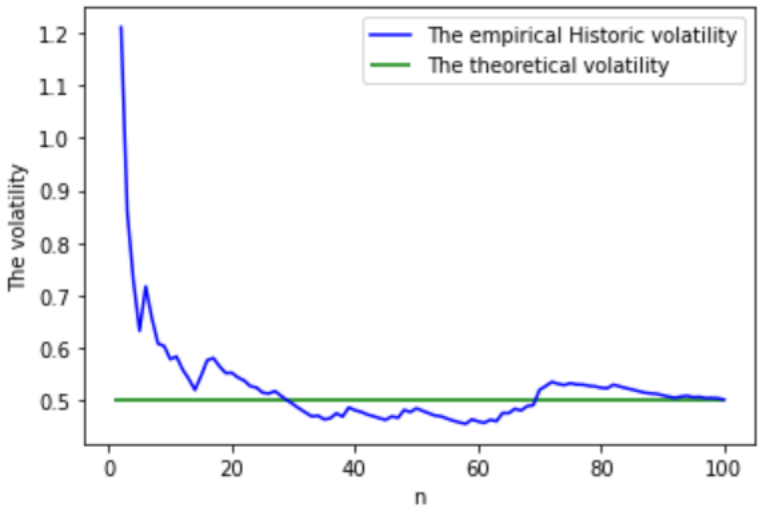
Historic volatility [4]: A statistical approach that consists of estimating the volatility as a standard deviation of the log-returns. This approach is intuitive because as we can see, the returns follow under the Black-Scholes model a log-normal distribution. $ln(\frac{S_t}{S_s}) \sim \mathcal{N}(0,(t-s)\sigma^2)$ for $t < s$. On that note, given a maturity T and a $\Delta t = \frac{T}{n}$ partition $0 = t_0 < t_1 < ... < t_n = T$ where $t_i = i.\frac{T}{n}$, $\forall i \in \{0,n\}$. We can easily see that the estimator $\tilde{\sigma}$ is consistent:

$$\hat{\sigma}^2 = \frac{1}{(n-1).\Delta t}.\sum_{i=1}^{n}(ln(\frac{S_{t_i}}{S_{t_{i-1}}}) - \bar{s})^2$$

Where $\bar{s}$ is the empirical mean : $\frac{1}{n}\sum_{i=1}^{n} ln(\frac{S_{t_i}}{S_{t_{i-1}}})$
In the following plot, we simulated GBM paths with a volatility parameter

$\sigma = 0.5$ and we tried to find it through the historic approach. In this case, we have an interesting estimation for $n = 100$

Figure 1: The historic volatility



Implied volatility : A computational approach based on empirical option prices. We retrieve this implied volatility by adjusting the volatility parameter in the *Black-Scholes* formula to replicate the empirical data.

$$\mathcal{C}(T, t, K, S_t, \sigma^{implied}) = \mathcal{C}^{observed}$$

This approximation is used for instance in the VIX index by averaging the implied volatility over the S&P-500 stock prices.

We will see as well that it is useful in pricing options under more sophisticated models, as well as describing the volatility surfaces under the *Dupire* model.

2.1 The time dependent volatility :

From the VIX index we can definitely see that in practice the volatility is not constant. In fact, the implied volatility surfaces under these assumptions are

flat, which is not consistent with the empirical volatility.
Let's consider now a time dependent model;

$$dS_t = S_t.\mu.dt + S_t.\sigma(t).dW_t^{\mathbf{P}}$$

To derive the option pricing formula, we can either apply the previous approach, or in practice use the *Dupire* equation [3] to find the implied *Black-Scholes* volatility (under the assumptions of *no-smile*):

$$\mathcal{C}(T, K, S_0, (\sigma_t)_{0\leq t\leq T}) = \mathcal{C}^{BS}(T, K, S_0, \sigma^{implied})$$

$$\begin{aligned}\sigma(T,K)^2 &= 2.\frac{\frac{\partial C^{BS}}{\partial T} + rK.\frac{\partial C^{BS}}{\partial K}}{K^2.\frac{\partial^2 C^{BS}}{\partial K^2}} \\ &= (\sigma^{implied})^2 + 2.\sigma^{implied}.T.\frac{\partial \sigma^{implied}}{\partial T} \\ \implies \sigma^{implied} &= \sqrt{\frac{\int_0^T \sigma(t)^2 dt}{T}}\end{aligned}$$

The no-smile assumption leads to a strike non dependent implied volatility. We can now use the *Black-Scholes* formula with the implied volatility to price a European call option under the time dependent model.

The historic volatility remains efficient in a time dependent model because we know that the returns are still governed by a log-normal distribution.

We generated some paths of the geometric Brownian motion where $S_0 = 100,\ T = 2,\ \sigma(t) = 0.2 + 0.1\ .\ sin(50.t)$:

We run a rolling standard deviation to find the historic volatility. Even though the results are not satisfyingly accurate because of the approximation errors both in simulating the stock prices and the window of the standard deviation, we can still find a sinusoidal distribution:

Figure 2: The GBM

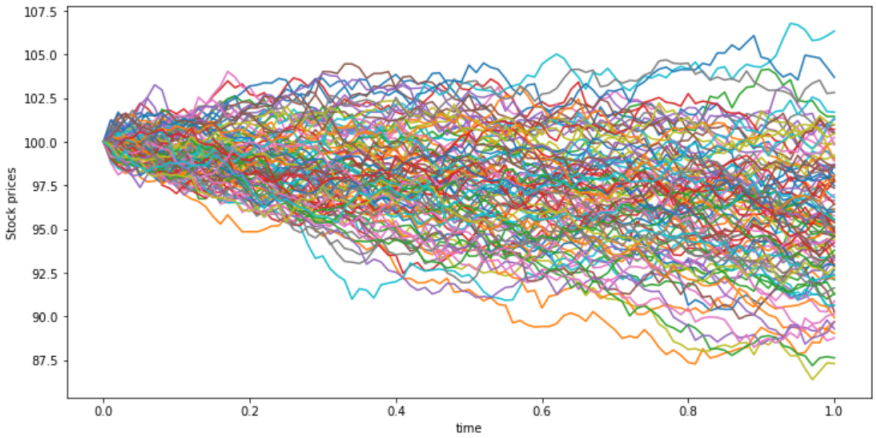
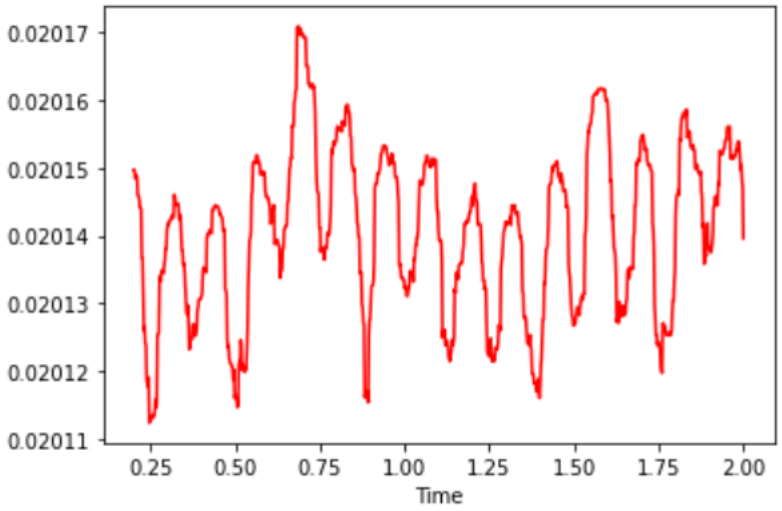


Figure 3: The historic volatility



3
LOCAL VOLATILITY

In the previous models, we were successful in establishing a statistical approach of the volatility. Yet, they do not replicate the empirical behaviors discussed above to the best. We have flat volatility surfaces, the correlation between volatility and stock prices is not present, and the most important one is the randomness that one can notice in the *VIX* index.

We need a source of randomness that could capture more characteristics and at the same time a source that could be tradable. This hypothesis is fundamentally important since it ensures the completeness of the market, which is essential to the further developments that we need.

Hence, the most intuitive source of randomness is the actual stock price. Thus, we write :

$$\frac{dS_t}{S_t} = \mu dt + \sigma(t, S_t).dW_t$$

we will see throughout the article that there is some kind of a trade off between the generality of the model and its robustness. Let's start with a common and simple calibration method in the *Dupire* model. We consider the total implied variance surfaces and the local volatility surfaces.

Using the *Dupire* equation we can find the relationship between the local volatility and the total implied variance $w(T, y) = T.(\sigma^{implied}(T, K))^2$ in terms if the moneyness [6] $y = ln(\frac{S_0}{K}) + \mu.T$

$$\sigma(T, K)^2 = \frac{\frac{\partial w}{\partial T}}{1 - \frac{y}{w}.\frac{\partial w}{\partial y} + \frac{1}{4}.(-\frac{1}{4} + \frac{1}{y} + \frac{y^2}{w^2}).(\frac{\partial w}{\partial y})^2 + \frac{1}{2}.\frac{\partial^2 w}{\partial y^2}}$$

In practice, we don't have access to these derivatives which can be problematic when reconstructing the volatility surfaces. As a first idea, we can use a finite difference method to approximate these derivatives. To approximate the

empirical values of the total implied variance, we can use a *Newton-Raphson* algorithm to find a zero of the equation $(**)$ for a fixed strike and maturity:

$$\mathcal{C}^{BS}(\sigma^{implied}) = \mathcal{C}^{observed} \ (**)$$

Let's consider now a data-set of European call option prices, with equidistant strikes and maturities $(\mathcal{C}_{i,j})_{1\leq i,j\leq N}$, $(T_i)_{1\leq i\leq N} = i.\Delta T$ and $(K_j)_{1\leq j\leq N} = j.\Delta K$. And let $\sigma_{i,j}^{implied}$ a solution for $(**)$ for (T_i, K_j).

$$\begin{cases} (\frac{\partial w}{\partial T})_{i,j} = \frac{(\sigma_{i+1,j}^{implied})^2.T_{i+1} - (\sigma_{i,j}^{implied})^2.T_i}{\Delta T} & 1 \leq i \leq N-1, 1 \leq j \leq N \\ (\frac{\partial w}{\partial K})_{i,j} = \frac{(\sigma_{i,j+1}^{implied})^2.T_i - (\sigma_{i,j}^{implied})^2.T_i}{\Delta K} & 1 \leq i \leq N, 1 \leq j \leq N-1 \\ (\frac{\partial^2 w}{\partial K^2})_{i,j} = \frac{(\sigma_{i,j+2}^{implied})^2.T_i + (\sigma_{i,j}^{implied})^2.T_i - 2.(\sigma_{i,j+1}^{implied})^2.T_i}{(\Delta K)^2} & 1 \leq i \leq N, 1 \leq j \leq N-2 \end{cases}$$

And,

$$\begin{cases} (\frac{\partial w}{\partial y})_{i,j} = (\frac{\partial w}{\partial K})_{i,j}.(-K_j) \\ (\frac{\partial^2 w}{\partial y^2})_{i,j} = \frac{(\sigma_{i,j+2}^{implied})^2.T_i.K_{j+1} + (\sigma_{i,j}^{implied})^2.T_i.K_j - (K_j + K_{j+1})(\sigma_{i,j+1}^{implied})^2.T_i}{(\Delta K)^2}.K_j \end{cases}$$

We can this way, approximate the local volatility surface. This method is naive, because in order to have better approximations for the derivatives we have to use an interpolation method like the linear or the spline interpolation.

We run the previous algorithm on European call option prices data for *Microsoft* (MSFT) stocks. The implied volatility surface (*Figure 4*) that we found is not flat, which is a good sign that the algorithm runs correctly. On the other hand, since the approximation of the local volatility (*Figure 5*) introduces derivatives that reduce significantly the efficiency of the algorithm, we found surfaces that are less smooth.

One important problem in this model is the robustness and the stability of the solutions. The finite difference method performs poorly as the grid strike-maturity changes even slightly, which makes it difficult to generalize the results of the standard *Dupire* model.

Figure 4: The implied volatility

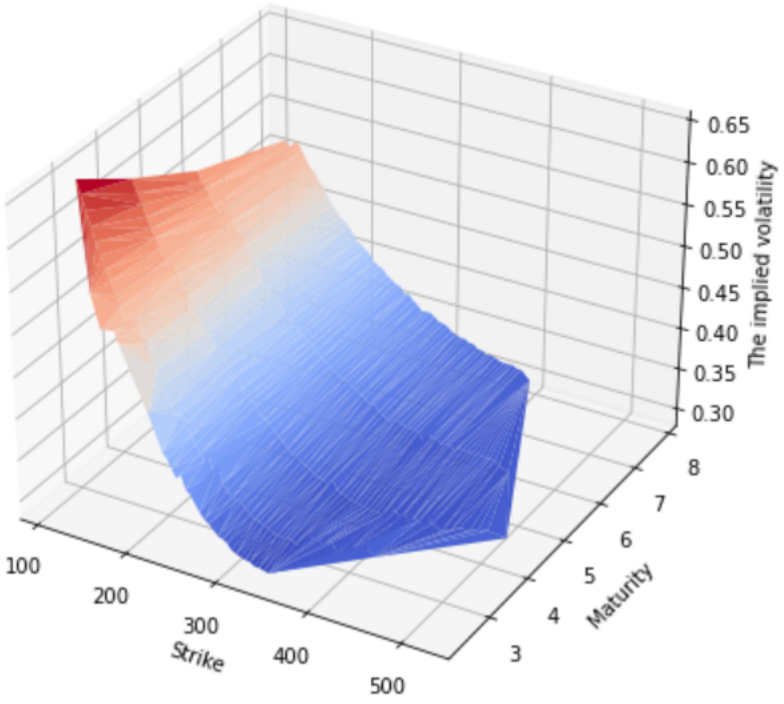
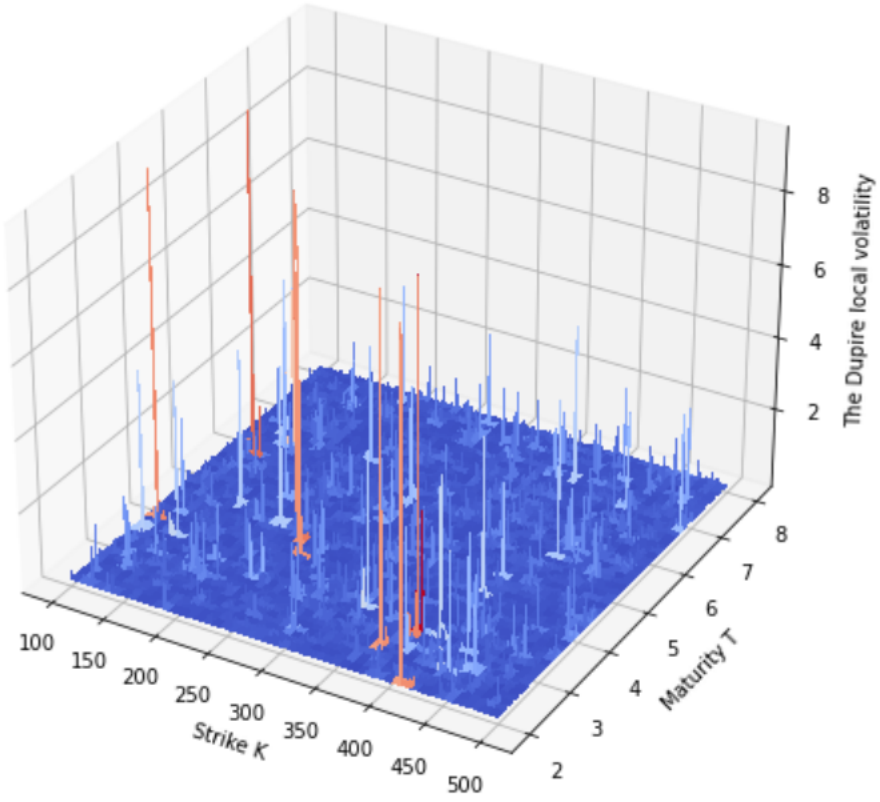


Figure 5: The local volatility



3.1 The *Constant Elasticity Variance*

In this model, the volatility is a power-law function of the level of the underlying asset.

$$\frac{dS_t}{S_t} = \mu dt + a.S_t^{n-1} dW_t,\ a > 0 \ and\ 0 \leq n \leq 1$$

And the volatility in this model is: $\sigma(t, S_t) = a.S_t^{n-1}$. This model is very interesting because for n=1 we find the *Black-Scholes* model, and for n=0 we find the *Bachelier* model.

We can capture more empirical features in this model, given that for n<1 we see the *leverage effect* and the *asymmetric volatility*.

Even though it is possible to computationally calibrate the *CEV* model using the formula of the European call option price at t for a maturity T and a strike K, for which we have an expression due to *Cox* for $n < 1$:

$$\mathcal{C}(\mathcal{S}, \tau) = S_t.\sum_{i=0}^{\infty} f_{1,i+1}(x).\Gamma(i+1+\frac{1}{2.(1-n)}, k.K^{2.(1-n)}) -$$
$$\text{K.e}^{-r.\tau} \sum_{i=0}^{\infty} f_{1,i+1+\frac{1}{2.(1-n)}}(x).\Gamma(i+1, k.K^{2.(1-n)})$$

Where :

$$\begin{cases} k = \frac{2\mu}{2.(1-n).a^2.(e^{2.(1-n).\mu.\tau}-1)}\ ,\ \tau = T-t \\ y = k.S_T^{2.(1-n)}\ ,\ x = k.S_t^{2.(1-n)}.e^{2.(1-n).\mu.\tau} \\ f_{l,m}(y) = \frac{l^m x^{m-1}.e^{-l.x}}{\Gamma(m)}\ (the\ Gamma\ density\ function) \end{cases}$$

But such formula is overwhelming and too heavy for calibration, we will instead examine a statistical approach of estimating the parameters of the *CEV* model from empirical stock prices.

Let's consider now an estimation method due to *Beckers* [1] for the parameters in the *CEV* model:

The first step in this method is to find a relationship between the instantaneous standard deviation of the returns, and the stock prices.

$$\begin{aligned} ln(std_t(\frac{dS_t}{S_t})) &= ln(a.S_t^{n-1}) \\ &= ln(a) + (n-1).ln(S_t) \end{aligned}$$

std_t is instantaneous and so it is the conditional standard deviation $std(.|S_t)$.

Now, let's suppose this linear equation is still for a finite period of time and not just instantaneous.

$$ln(std_t(\frac{S_{t+1}}{S_t})) = \alpha + \beta.ln(S_t) \ (*)$$

The only difference this time is that we do not know the expression of α and β in terms of a and n.

The general problem is very complicated, because we don't know what α and β stand for concretely. On that note, we will consider the square-root model as a particular case (where n=$\frac{1}{2}$), in which we can understand why this assumption is valid. We admit the following result, that is due to *Feller* [5]: In the square-root model the conditional distribution of S_T on S_t can be written as

$$p_t(S_T) = kx. \sum_{n=0}^{\infty} \frac{e^{-x}.x^n}{(n+1)!}.f_{1,n+1}(y)$$

Where:

$$\begin{cases} k = \frac{2\mu}{a^2.(e^{\mu.\tau}-1)} \ , \ \tau = T-t \\ y = k.S_T \ , \ x = k.S_t.e^{\mu.\tau} \\ f_{l,m}(y) = \frac{l^m x^{m-1}.e^{-lx}}{\Gamma(m)} \ (the \ Gamma \ density \ function) \end{cases}$$

The moments of S_T can be derived from the previous formula:

$$\mathbb{E}_t(S_T^j) = \frac{x}{k^j}. \sum_{n=0}^{\infty} \frac{e^{-x}.x^n.(n+j)!}{n!.(n+1)!}$$

In particular, we calculate $\mathbb{E}_t(S_T) = \frac{x}{k}.\sum_{n=0}^{\infty}\frac{e^{-x}.x^n.(n+1)!}{n!.(n+1)!} = \frac{x}{k}$, and $\mathbb{E}_t(S_T^2) = \frac{x}{k^2}.\sum_{n=0}^{\infty}\frac{e^{-x}.x^n.(n+2)}{n!} = \frac{x}{k^2}.(x+2)$

Hence, the conditional variance is:

$$\begin{aligned}
\mathbb{V}ar_t(\frac{S_T}{S_t}) &= \frac{\mathbb{V}ar_t(S_T)}{S_t^2} \\
&= \frac{\frac{2.x}{k^2}}{S_t} \\
&= \frac{2.S_t.e^{\mu.\tau}}{k.S_t} \ (the\ expression\ of\ x\ and\ k) \\
&= \frac{e^{\mu.\tau}}{\mu.S_t}.a^2.(e^{\mu.\tau}-1)
\end{aligned}$$

Let's substitute in the previous computations with $\tau = 1$.

$$std_t(\frac{S_{t+1}}{S_t}) = \sqrt{\frac{e^{\mu}}{\mu.S_t}.a^2.(e^{\mu}-1)}$$

$$\implies ln(std_t(\frac{S_{t+1}}{S_t})) = -\frac{1}{2}.ln(S_t) + \frac{1}{2}.ln(\frac{e^{\mu}}{\mu}.a^2.(e^{\mu}-1))$$

This result is interesting because we can both prove (*) and find α and β in terms of the *CEV* parameters.

From here, we can run a linear regression on (*) in order to find an estimation of the constant volatility a, and to validate the *CEV* model. (We find out how much the power is different than $-\frac{1}{2}$)

We run the previous approach on 50 stocks tracked in the S&P 500 index from April 2020 to April 2022. To find the instantaneous standard deviation we run a rolling standard deviation on the stock prices. Then we estimated the slope and the intercept of the linear regression. The histogram shows the distribution of the slope and the intercept amongst the tracked stocks.

Figure 6: The slope

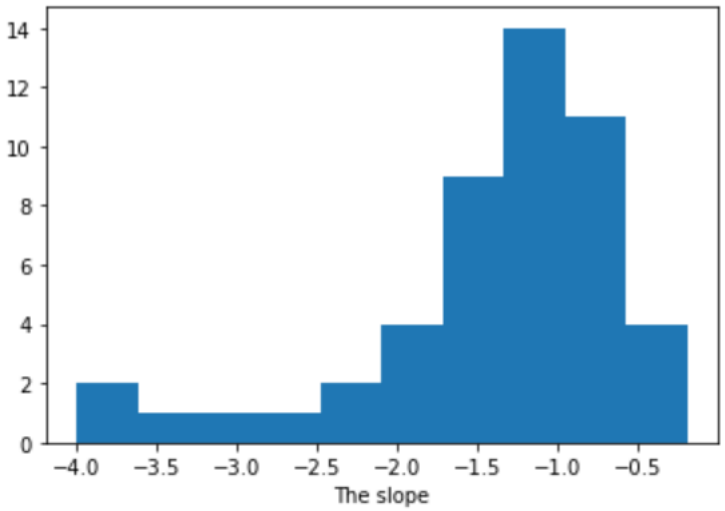
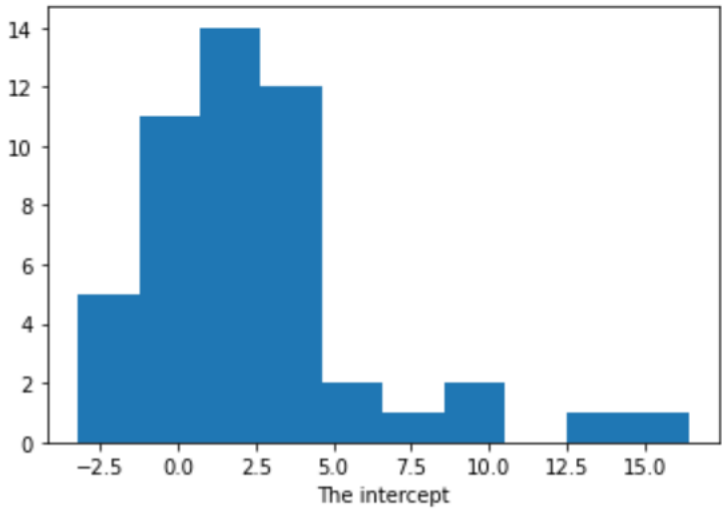


Figure 7: The intercept



An important observation about these results is the slope which validates the *CEV* and the asymmetric volatility effect. We can see that the slopes are more distributed around -1 which somehow brings us back to the *Bachelier* model.

Since the inverse relationship between stock price and volatility is supported by the data, a quick check can be performed to verify the underlying rationale. Fluctuations in the debt-equity ratio due to changes in the stock price are thought to cause the volatility to move in the opposite direction of the stock price.

In the particular case of the square-root model, we can approximate the constant volatility parameter without having to use a linear regression or even a standard deviation. The problem with the standard deviation approach is the insufficiency of the data and the size of the rolling method's window. It is usually difficult to adjust empirically the parameters to ensure a better estimation.

We saw earlier that :

$$\mathbb{V}ar_t(\frac{S_T}{S_t}) = \frac{e^{\mu.\tau}}{\mu.S_t}.a^2.(e^{\mu.\tau} - 1)$$

If we assume that the log-normal diffusion still holds

$$\mathbb{V}ar(\frac{S_T}{S_0}) = e^{2.T.\mu}.(e^{\sigma_{BS}^2 T} - 1)$$

$$\implies a^2 = \frac{\mu.S_0.e^{\mu.T}.(e^{\sigma_{BS}^2 T} - 1)}{e^{\mu.T} - 1}(with\ \tau = T)$$

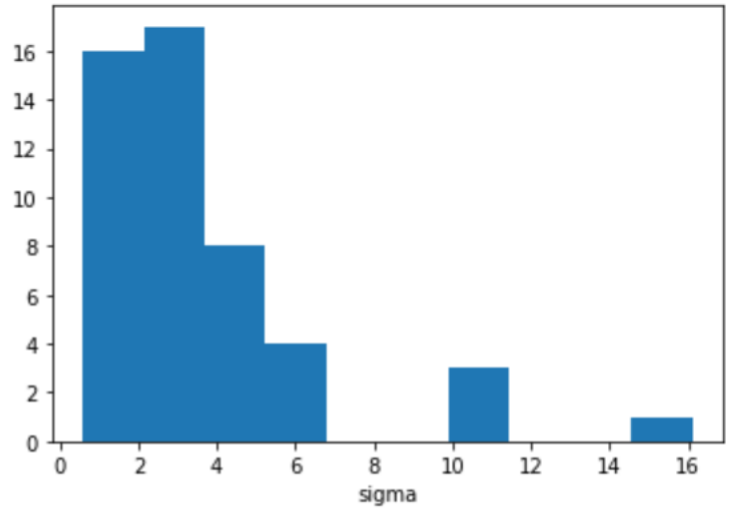
As for σ_{BS}, we can use the *historic* volatility to provide an estimate.

We can apply this approach to the previous data set with $e^{\mu.T} = \mathbb{E}(S_T)$. In practice, given only one sample of the stock price, we could split the data set into different sub-sets and run the mean over all these sub-sets (considering the sub-set different samples of the same distribution).

For instance in a data set of T = 1000, we have $e^{\mu.100} = \frac{1}{10}.\sum_{i=1}^{10} S_{100.i}$ and retrieve an approximation of the drift.

As we can see, there are some differences with the previous plot. But, overall the parameters is mostly distributed around 2, which is the case as well for the previous method.

Figure 8: The constant volatility



After estimating the parameters of the *CEV* model, we can see how the implied volatility surfaces look under these assumptions.

The previous *Dupire* equation can be seen as a PDE of the total implied variance.

$$\frac{\partial w}{\partial T} = (1 - \frac{y}{w}.\frac{\partial w}{\partial y} + \frac{1}{4}.(-\frac{1}{4} + \frac{1}{w} + \frac{y^2}{w^2}).(\frac{\partial w}{\partial y})^2 + \frac{1}{2}.\frac{\partial^2 w}{\partial y^2}).\sigma.K(y)^{n-1}$$

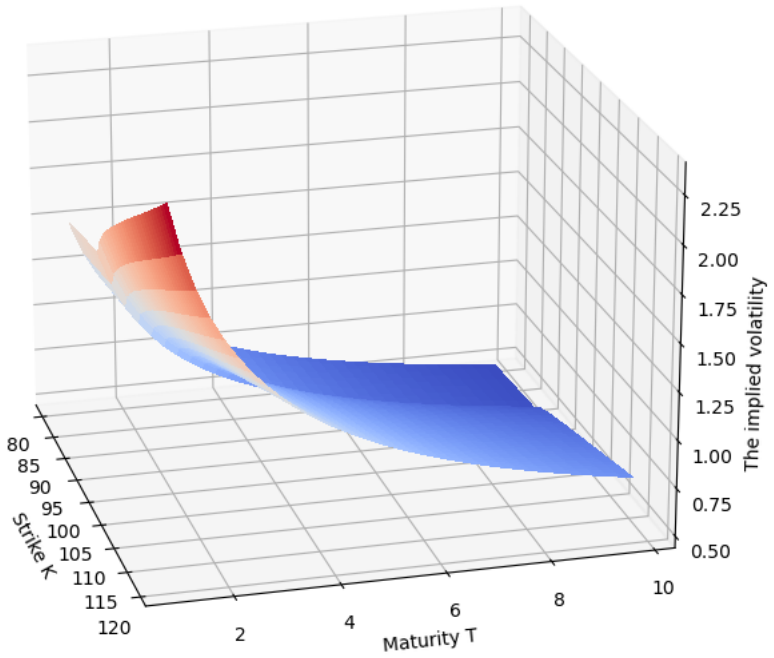
with, $w(0, y) = 0$ and $w(T, 0) = m$ (at the money value of the implied total variance).

Te visualize the results, we made some exhaustive assumptions :

$$\begin{cases} \frac{\partial w}{\partial T} \sim (\sigma^{implied}(T,K))^2 \\ w(T,y) \sim m + \frac{\partial w}{\partial y}.y \\ \frac{\partial^2 w}{\partial y^2} \sim 0 \end{cases}$$

In this case, we can see that the surface in *figure 9* is smooth and represents the volatility smile. The estimation parameters in the *CEV* model were taken for *Amazon.com Inc. (AMZN)* stock prices. Here we found the implied volatility surface that ensure a CEV local volatility.

Figure 9: The implied volatility



3.2 The mixture distribution model

The assumption of a constant volatility that should be used to price any derivative security with the same underlying asset fails to hold true in practice. The previous local volatility model is an interesting one, given that it can capture a wide range of empirical properties, such as the asymmetric volatility effect and

the volatility smile.

As mentioned above, empirical stock prices typically have skewed implied volatility shapes, *Damiano Brigo and Fabio Mercurio* [2] introduced a class of analytically tractable models for an asset-price dynamics that are flexible enough to recover a large variety of market volatility structures.

In the particular case of a log-normal mixture distribution, we can capture more empirical many empirical characteristics. Let's assume the risk-neutral probability density function $p(t, S)$, at a time t for a stock price S_t, to be a weighted sum of different log-normal probability densities $(p_i(t, S))_{1 \leq i \leq N}$.

$$p(t, S) = \sum_{i=1}^{N} w_i(t).p_i(t, S)$$

Assuming the distributions have the same mean μ and different variances $(\sigma_i(t))_{1 \leq i \leq N}$ we have,

$$p_i(t, S) = \frac{1}{S.\sigma_i(t).\sqrt{2\pi}}.exp(-\frac{(ln(S) - \mu)^2}{2.\sigma_i(t)^2})$$

Under the previous assumptions, the local volatility is:

$$\sigma^2(t, S) = \sum_{i=1}^{N} w_i(t).\sigma_i(t)^2$$

This approach is very interesting given the pricing formula of European call options:

$$\mathcal{C}(T, K, S_0, \sigma) = \sum_{i=1}^{N} w_i(t).\mathcal{C}_{time-dependent}(T, K, S_0, (\sigma_i(t))_{t \geq 0})$$

As mentioned above, in the time dependent model, we can price options using the implied *Black-Scholes* volatility. Therefore,

$$\begin{cases} \mathcal{C}(T, K, S_0, \sigma) = \sum_{i=1}^{N} w_i(t).\mathcal{C}^{BS}(T, K, S_0, \eta_i) \\ \eta_i = \sqrt{\frac{\int_0^T \sigma_i(t)^2 dt}{T}}, \ \forall i \in [1, N] \end{cases}$$

Unlike the *CEV* model, this approach is practical in pricing options and approximating the implied volatility surfaces.

Approximating the implied volatility $\hat{\sigma}(y)$, with $y := ln(\frac{S_0}{K}) + \mu.T$ is hence equivalent to solving the equation :

$$\mathcal{C}(T, K, S_0, \hat{\sigma}(y)) = \sum_{i=1}^{N} w_i(t).\mathcal{C}^{BS}(T, K, S_0, \eta_i) \ (***)$$

(***) can be reformulated as (assuming that the asset is driftless);

$$\phi(\frac{y + \frac{\hat{\sigma}(y)^2}{2}.T}{\hat{\sigma(y)}.\sqrt{T}}) - e^{-y}.\phi(\frac{y - \frac{\hat{\sigma}(y)^2}{2}.T}{\hat{\sigma(y)}.\sqrt{T}}) = \sum_{i=1}^{N} w_i(t) \left[\phi(\frac{y + \frac{\eta_i^2}{2}.T}{\eta_i.\sqrt{T}}) - e^{-y}.\phi(\frac{y - \frac{\eta_i^2}{2}.T}{\eta_i.\sqrt{T}})\right]$$

Where ϕ is the standard normal cumulative distribution function. *Brigo and Mercurio* derived from this equation an approximation of the implied volatility using a second order Taylor expansion of the function $\hat{\sigma}$ around 0. They computed the derivatives of $\hat{\sigma}$ through the repetition of *Dini*'s implicit function theorem.

$$\hat{\sigma}(y) = \hat{\sigma}(0) + \frac{1}{2.\hat{\sigma}(0).T}. \sum_{i=1}^{N} w_i(t). \left[\frac{\hat{\sigma}(0)}{\eta_i}.e^{\frac{1}{8}(\hat{\sigma}(0)^2 - \eta_i^2).T} - 1\right].y^2 + o(y^2)$$

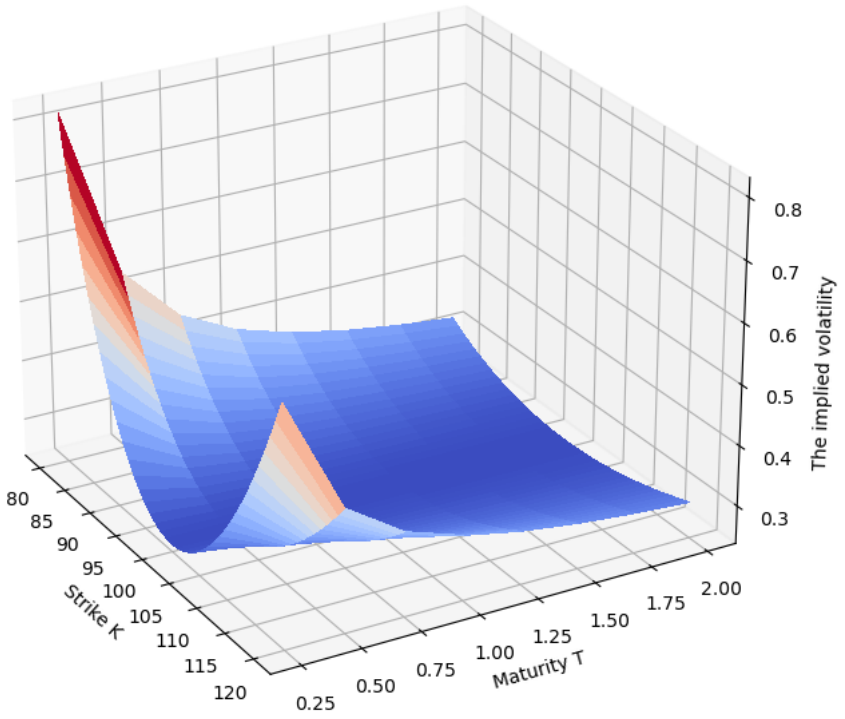
$\hat{\sigma}(0)$ is given using the previous equality for $y = 0$:

$$\phi(\frac{\hat{\sigma}(0).\sqrt{T}}{2}) - \phi(-\frac{\hat{\sigma}(0).\sqrt{T}}{2}) = \sum_{i=1}^{N} w_i(t) \left[\phi(\frac{\eta_i.\sqrt{T}}{2}) - \phi(-\frac{\eta_i.\sqrt{T}}{2})\right]$$

$$\implies 2.\phi(\frac{\hat{\sigma}(0).\sqrt{T}}{2}) - 1 = 2. \sum_{i=1}^{N} w_i(t) \left[\phi(\frac{\eta_i.\sqrt{T}}{2})\right] - 1$$

$$\implies \hat{\sigma}(0) = \frac{2}{\sqrt{T}}.\phi^{-1}\left(\sum_{i=1}^{N} w_i(t) \left[\phi(\frac{\eta_i.\sqrt{T}}{2})\right]\right)$$

Figure 10: The implied volatility surface for N=2, μ=0, r=0.05, $T \in \{0.25, 0.5, 0.75, 1, 1.25, 1.5, 1.75, 2\}$, $K \in \{80, ..., 120\}$, $S_0 = 100$, $(w_1, w_2) = (0.6, 0.4)$, $(\eta_1, \eta_2) = (0.35, 0.1)$)



4 STOCHASTIC VOLATILITY

4.1 Definition :

Although local volatility models were an improvement on time dependent volatility, they possessed certain undesirable properties. For example, volatility is perfectly correlated (positively or negatively) with stock price yet empirical observations suggest no perfect correlation exists. Stock prices empirically exhibit volatility clustering but under local volatility this does not necessarily occur.

Consequently after local volatility development, models were proposed that allowed volatility to be governed by its own stochastic process. We now define stochastic volatility.

The word "stochastic" means that some variable is randomly determined and cannot be predicted precisely. In stochastic volatility models, volatility has its own source of randomness making volatility intrinsically stochastic. We can therefore never definitely determine the volatility's value, unlike in local volatility.

Hence, the stochastic differential equation becomes:

$$dS_t = S_t.\mu.dt + S_t.\sigma(w).dW_t^{\mathbf{1}}$$

where the volatility $\sigma(w)$ is governed by its own stochastic process $dW_t^{\mathbf{2}}$. The Wiener processes have instantaneous correlation $\rho \in [-1, 1]$ defined by :

$$corr(dW_t^{\mathbf{1}}, dW_t^{\mathbf{2}}) = \rho.dt$$

Stochastic volatility models tend to be analytically less tractable. In fact, it is common for stochastic volatility models to have no closed form solutions for option prices. Consequently option prices can only be calculated by simulation. Thus, we will be seeing the Heston model that stands out from other stochastic volatility models because there exists an analytical solution for European options.

4.2 Heston model :

In the case of the Heston model [8], the stochastic differential equation becomes:

$$dS_t = r.S_t.dt + S_t.\sqrt{V_t}.dW_t^{\mathbf{1}}$$

$$dV_t = \kappa.(\theta - V_t).dt + \sigma.\sqrt{V_t.dt}.dW_t^{\mathbf{2}}$$

where:

- $W^{\mathbf{1}}$ and $W^{\mathbf{2}}$ are two Wiener processes and $corr(dW_t^{\mathbf{1}}, dW_t^{\mathbf{2}}) = \rho.dt$
- κ**:** Variance mean reversion speed
- θ**:** Long term variance
- $V0$**:** Initial variance
- σ**:** Volatility of variance
- ρ**:** Spot v.s. variance correlation

The Heston model allows for semi-analytical solutions to option pricing. The value of a European call option can be obtained by using a probabilistic approach:

$$C = S_0.\Pi_1 + K.e^{-rT}.\Pi_2$$

with

$$\Pi_1 = \frac{1}{2} + \frac{1}{\pi}.\int_0^{+\infty} Re\left(\frac{e^{-iwln(K)}.\psi(w-i)}{iw\psi(-i)}\right)\,dw$$

$$\Pi_2 = \frac{1}{2} + \frac{1}{\pi}.\int_0^{+\infty} Re\left(\frac{e^{-iwln(K)}.\psi(w)}{iw}\right)\,dw$$

where ϕ is the characteristic function of the log-price: $\phi(w) = \mathbb{E}\left[e^{iwln(S_T)}\right]$

Hence, the semi-analytical formula for a Heston call is :

$$C = \frac{1}{2}(S_0 - Ke^{-rT}) + \frac{1}{\pi}.Re\left[\int_0^{+\infty}\left(\frac{e^{rT}.\phi(w-i)}{iwK^{iw}} - \frac{K.\phi(w)}{iwK^{iw}}\right)\,dw\right]$$

with :

$$\phi(u) = e^{rT}(S_0)^{iu}\left(\frac{1-ge^{-dT}}{1-g}\right)^{-2\frac{\theta\kappa}{\lambda^2}} \exp\left(\frac{\theta\kappa T}{\lambda^2}.(\kappa - i\rho.\lambda u - d) + \frac{V_0}{\lambda^2}.(\kappa - i\rho.\lambda u + d).\frac{1-e^{-dT}}{1-ge^{-dT}}\right)$$

- $d = \sqrt{(\rho\lambda ui)^2 + \lambda^2(ui + u^2)}$
- $g = \frac{\kappa - \rho\lambda ui - d}{\kappa - \rho\lambda ui + d}$
- λ is the the market price of volatility risk

In the following plot, we generated some paths of stock prices and volatilities with Heston model where $T = 1, S_0 = 55, K = 50, r = 0.04, V_0 = 0.04, \rho = -0.7, \kappa = 2, \theta = 0.04$ and $\sigma = 0.3$:

Figure 11: Stock prices

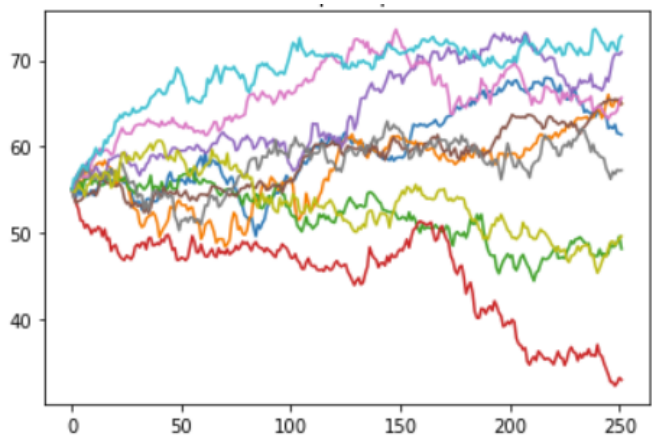
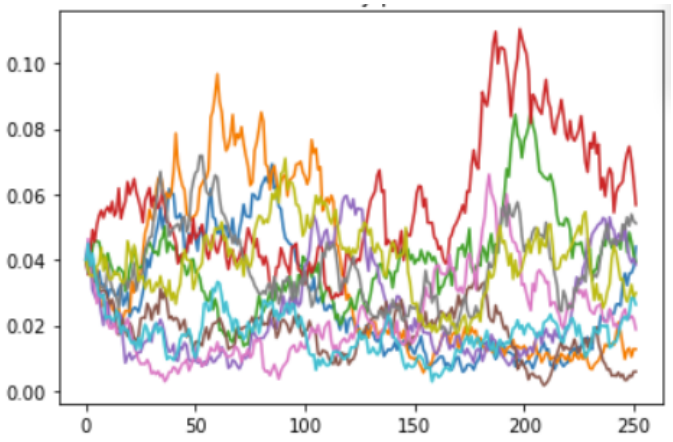


Figure 12: Volatilities



4.3 Calibration

Calibration of the parameters $\Theta = (V_0, \kappa, \theta, \sigma, \rho, \lambda)$ is an important long standing problem in quantitative finance [9]. Not only a solution must be found which reproduces the data but it must be stable in the sense that adding new data should not change the parameters drastically.

Let's consider the Heston model parameterized by a set of parameters $\Theta = (V_0, \kappa, \theta, \sigma, \rho, \lambda)$.

We will analyze call options characterized by strike K and maturity T. The market pricing function P^{MKT} then takes two parameters K and T and outputs market prices of options characterized by their strike and maturity, which we denote by $P^{MKT}(T, K)$ for each T and K.

The Heston model pricing function P^H then takes three parameters Θ, K and T. It outputs prices of options following the Heston model characterized by their strike, maturity and the set of parameters Θ. Thus, we denote it by $P^H(\Theta, T, K)$ for each Θ, T and K.

The common property of all model calibration methods is that they iteratively evaluate the model pricing function, or some approximation of it, on each instance of model parameters Θ until a small enough distance, described by some appropriate distance function, between model prices and market prices is obtained. Formally, this can be described by the minimization :

$$\hat{\Theta} = \underset{\Theta}{\operatorname{argmin}}\, d\left(P^H(\Theta, T, K), P^{MKT}(T, K)\right)$$

where d is some distance function.

One can let the function d in Equation above be related to the squared distance, then the calibration becomes:

$$\hat{\Theta} = \underset{\Theta}{\operatorname{argmin}} \sum_{i=1}^{N} \sum_{j=1}^{M} w_{i_j} \left(P^H(\Theta, T_i, K_j) - P^{MKT}(T_i, K_j)\right)^2$$

where the dimensionality of the grid is given by N×M, that is, N maturities and M strikes, and the weights w is pre-specified and can reflect the relative

importance of a specific option.

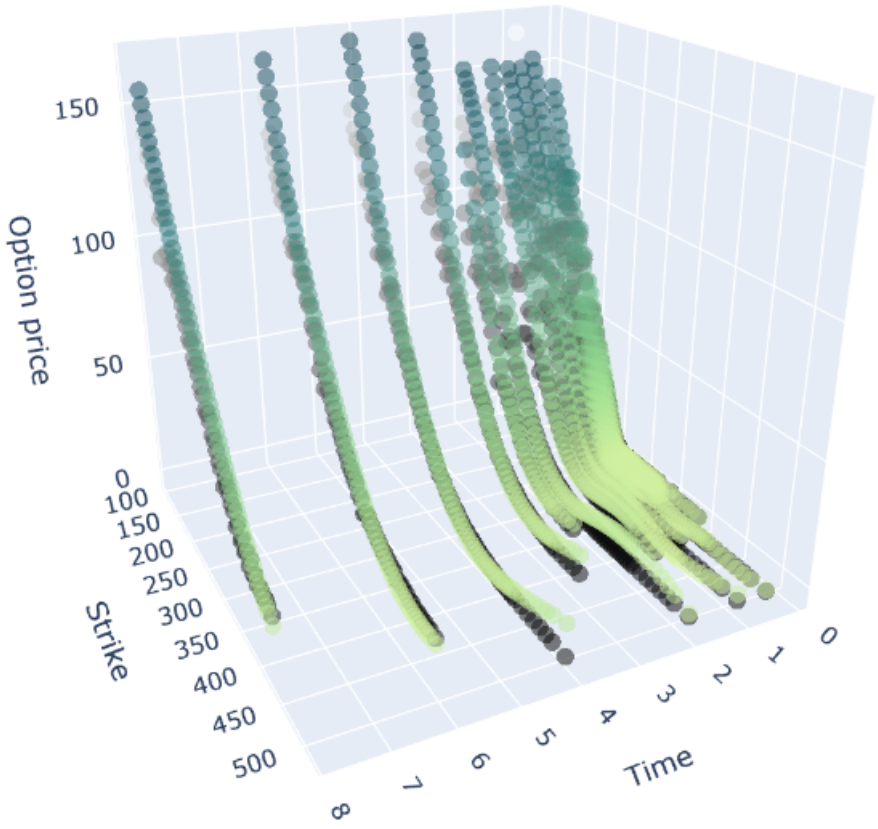
In the simulation below, we will calibrate a Heston model using Microsoft call options by minimizing the squared distance between model and market prices. We suppose that all options have the same importance, thus, our weights are all equal to 1.
Boundaries are specified in order to limit the optimization to a reasonable set of Heston model parameters:

Parameter	Boundaries
V_0	$[0.001, 0.1]$
κ	$[0.001, 5]$
θ	$[0.001, 0.1]$
σ	$[0.01, 1]$
ρ	$[-1, 0]$
λ	$[-1, 1]$

As a result, We find our optimal parameters for the calibrated Heston model and we compare the prices following this model with market prices in the figure below:

Parameter	Value
V_0	0.001
κ	5
θ	0.020
σ	0.251
ρ	-0.999
λ	1

Figure 13: Market prices (gray) vs Heston model prices (green)



5 DISCUSSION

We have surveyed the key volatility models and developments, highlighting the innovation associated with each new class of volatility models. In conclusion it can be seen from our review that these developments have progressed in a logical order to address key shortcomings and anomalies of previous models.

Time dependent models addressed option prices varying with expiration dates, local volatility also addressed volatility smiles and the leverage effect, whereas stochastic volatility could incorporate all the effects captured by local volatility and a range of other empirical effects such as a greater variability in observed volatility along with the mean reversion effect. The next step would be to take up more advanced stochastic volatility models, which can take fully into account the strong auto-regressive behaviors tracked in practice. The discrete time GARCH(p,q) process for the volatility would be an alternative. And we can pursue even further studies oriented towards the jump diffusion models due to their consistency with the empirical stock prices.

However the trade-off associated with improved volatility modeling has been at the expense of analytical tractability.

REFERENCES

[1] Stan Beckers. "The Constant Elasticity of Variance Model and Its Implications For Option Pricing". In: *THE JOURNAL OF FINANCE* (1980).

[2] Brigo and Mercurio. "Lognormal-mixture dynamics and calibration to market volatility smiles". In: (2002).

[3] B. Dupire. "Pricing with a smile". In: *Risk 7* (1994).

[4] Louis H. Ederington. "Measuring Historical Volatility". In: *JOURNAL OF APPLIED FINANCE* (2006). DOI: `https://papers.ssrn.com/sol3/papers.cfm?abstract_id=926167`.

[5] W. Feller. "An Introduction to Probability Theory and its Application". In: *THE JOURNAL OF FINANCE* (1979).

[6] Jim Gatheral. *The Volatility Surface: A Practitioner's Guide*. Wiley; 1st edition. Wiley, 2006.

[7] Sovan Mitra. "A Review of Volatility and Option Pricing". In: *Imperial College* (2009). DOI: `https://arxiv.org/pdf/0904.1292.pdf`.

[8] Olivier Pironneau. "Calibration of Heston Model with Keras". In: (2019).

[9] Oliver Klingberg Malmer Victor Tisell. "Deep Learning and the Heston Model: Calibration Hedging". In: (2020).