City, University of London Institutional Repository

This is the published version of the paper.

This version of the publication may differ from the final published version.

Monotonic transformation and recovering the implied stock price process

Gianluca Fusai[1,2]

Abstract

The paper demonstrates the construction of a stock price stochastic process that aligns with observed option prices, drawing inspiration from a manuscript by Professor Erio Castagnoli that employs the concept of a non-linear monotonic transformation of a standard Brownian motion. The paper includes a detailed numerical example that illustrates the implementation of this straightforward and ingenious idea.

Keywords Brownian motion · Implied volatility surface · Monte Carlo simulation · Black-Scholes model

Mathematics Subject Classification 35K99 · 65D05

JEL Classification C02 · C63 · G13 · G17

1 Preface

For several months, I had been wondering what to write in remembrance of Erio. Until, during a rainy Sunday afternoon that I was spending cluttering my study, I came across a folder containing a short manuscript by Erio on g-normal transformations, see Castagnoli (2001).[1] Surprisingly, I had completely forgotten about it, but now I vividly recall the meeting we had. I remember that Erio had given me these notes because, at that time, my friends Emanuele Amerio, Antonio Vulcano, and I were attempting to model the risk-neutral dynamics of the whole volatility surface, see Amerio et al. (2003). The implied volatility surface is a graphical representation

---

✉ Gianluca Fusai
  gianluca.fusai.1@city.ac.uk; gianluca.fusai@uniupo.it

[1] Faculty of Finance, Bayes Business School (formerly Cass), City, University of London, 106 Bunhill Row, London EC1Y 8TZ, UK

[2] Department SEI, Università del Piemonte Orientale, Via Perrone 18, 28100 Novara, Italy

[1] A copy of the notes is available to those who desire them.

depicting implied volatilities across various option strikes and expirations. Implied volatility serves as a gauge for the market's anticipation of future price volatility in an underlying asset, inferred from current option prices. For a given maturity, the implied volatility curve can exhibit different shapes for distinct strike prices, such as a smile, smirk, or tilt. A smile occurs when implied volatilities are elevated for both out-of-the-money (OTM) and in-the-money (ITM) options compared to at-the-money (ATM) options. Conversely, a smirk or tilt arises when the skew is more pronounced on one side, either higher for OTM or ITM options. The presence of a non-flat implied volatility curve signals that the stock price distribution implied in option prices deviates from the lognormal distribution supporting Black-Scholes option pricing.

During the meeting, which I can approximately date back to 2001, Erio proposed a novel approach involving the modeling of a hidden factor, potentially using an Arithmetic Brownian Motion (ABM). The objective was to attain the desired stock price distribution through a non-linear, monotonic transformation. Erio called it g-transformation.[2]

At the conclusion of the meeting, Erio kindly made photocopies of his notes for me. Only now, upon rediscovering these copies, I am struck by their simplicity yet profound insights. The core idea is to identify a diffusion process for the stock price that, at specific option maturities, exhibits a distribution in line with the observed option prices. This is achieved by modeling stock returns as a monotonic generalized transformation, hence the term g-normal, of an arithmetic Brownian motion (ABM). With the function g in place, deriving the dynamics of the log-returns became relatively straightforward through the application of Ito's lemma. These precise steps were outlined in Erio's handwritten notes.

In fact, if this g were linear, also the log-returns would have a normal distribution and for example the Black-Scholes formula for option pricing holds. So to obtain non-normal distributions it is necessary to choose g in a suitable and general way. The name g-normal therefore also came naturally, although it could also generate some confusion if one thinks of the log-normal. In fact, here the aim is not to move from the distribution of returns (assumed to be normal) to that of prices (that therefore have a log-normal distribution), but rather to try to figure out how to model returns as a non-linear g-function of a hidden process modelled according to a Brownian motion.

---

[2] The idea of studying monotonic transformation of normal r.v.'s was not new. Indeed, Erio had previously recommended that I investigate a similar idea between 1990–92. In particular, He suggested exploring the relationship between mean-variance (MV) and second order stochastic dominance (SSD) in the context of transforming normal random variables. The specific question he recommended investigating was: What conclusions can be drawn about the stochastic dominance relationship between two monotonic transformations of normal random variables, namely, $Y_1 = g(X_1)$ and $Y_2 = g(X_2)$? It is well-known that MV dominance and SSD produce equivalent orderings for normal random variables. However, this equivalence is not guaranteed when examining stochastic dominance orderings between non-normal random variables. Interestingly, it turns out that the quantities involved in establishing an equivalence between MV and SSD, are not the mean and variance of Y, but rather the mean of Y and the variance of X. This can be easily demonstrated, for example, by considering the lognormal distribution, where g is the exponential function. This idea served as inspiration for my first paper, which I presented at the AMASES Conference in Treviso in 1992, see Fusai (1992). I am still deeply grateful to Francesca Beccacece and Margherita Cigola for their invaluable assistance in revising my first paper and preparing the presentation I delivered at the AMASES conference. Modesti (1993) had also studied a similar problem.

As I turned to the second page of the manuscript, Erio's note presented a straightforward formula to determine the desired function g, as detailed in Eq. (1) in the next section. Notably, a one-to-one monotonic transformation can always be identified between the specified distribution, such as the one derived from option prices, and the normal distribution. The function g precisely embodies this transformation, facilitating the conversion between the hidden process and the risk-neutral stock price process. Furthermore, employing a standard Brownian motion (SBM) process offers distinct advantages. By combining the g-transformation with SBM dynamics, it becomes feasible, through a simple application of Ito's lemma, to generate and then simulate log-return process dynamics that align seamlessly with the specified distribution extracted from option prices.[3]

In conclusion, the intuition in Erio's note is that, the knowledge of the function g facilitates the development of a stochastic dynamics model for the stock return: I just need to simulate a BM and then apply to it, at any point in time, the function g. The additional contribution of this paper is to build on the Erio's intuition and to put it in practice in the context of recovering the stock price process implied in option prices.

The paper is structured as follows. Section 2 formalizes the idea. Section 3 presents the construction of the cumulative distribution from option prices. Section 4 presents an empirical application. Section 5 illustrates possible application and extensions. Then I conclude.

2 The *g*-normal process

Let $\{X_t\}_{t \geq 0}$ be a SBM process so that at time t

$$X(t)|\mathcal{F}_0 \sim \mathcal{N}(0, t),$$

where $\mathcal{F}_t$ is the filtration generated by the Brownian motion.

Following Erio's notes, I assume that there exists some (non-linear) mapping between $X(t)$ and the stock log-return

$$r(0, t) = \ln(S(t)/S(0)) = g(X(t), t).$$

Equivalently,

$$S(t) = S(0)e^{g(X(t),t)}.$$

---

[3] The idea is clearly not new. For example, an important article by Mark Rubinstein on Implied Trees, see Rubinstein (1994), presents a method for constructing a binomial tree that is consistent with observed option prices. The implied binomial tree can be used to price a wide range of options, including ones with early exercise features. The method is based on the principle of no-arbitrage and involves iteratively adjusting the binomial tree until the option prices implied by the tree match the observed option prices. The implied binomial tree has become a widely used tool in option pricing and has been extended and adapted in many subsequent papers.

Table 1 Moving from a SBM to the log-return with CDF F

Random variable	$X(t)$	$Z(t) = \frac{X(t)}{\sqrt{t}}$	$U(t) = \Phi\left(\frac{X(t)}{\sqrt{t}}\right)$	$r(0, t) = F^{-1}\left(\Phi\left(\frac{X(t)}{\sqrt{t}}\right); t\right)$
Distribution	$\mathcal{N}(0, t)$	$\mathcal{N}(0, 1)$	$\mathcal{U}(0, 1)$	$F(r; t)$

In addition, let F be the cumulative distribution function (CDF) of log-return

$$F(x; t) = \mathbb{P}_0(r(0, t) \leq x).$$

If the function g is linear, then the returns have a normal distribution, the stock price is lognormal and the Black-Scholes model for option pricing holds. Far from g being linear, far the distribution of r from being normal, the stock price from having a lognormal distribution, and the Black-Scholes formula for options to hold.

The first remark concerns the determination function g. The result is located on page 3 of the manuscript.[4] It can be written in terms of the standard normal cumulative distribution and the inverse cumulative distribution of F

$$g(X(t), t) = F^{-1}\left(\Phi\left(\frac{X(t)}{\sqrt{t}}\right); t\right) \tag{1}$$

where F^{-1} is the inverse of the stock return distribution. Formula (1) has a straightforward interpretation. Indeed, the r.v. $U = \Phi\left(\frac{X}{\sqrt{t}}\right)$ has a uniform distribution and then $F^{-1}(U; t)$ has the assigned CDF F. The transformation described above is indeed one of the well-known approaches for simulating a random variable given its distribution function. For example, Monte Carlo simulations of a normal random variable, starting from a pseudo-uniform distribution, benefit from these steps. Notice that the equivalence between the r.v. having the distribution implied by option prices and the transformed SBM is valid only in law and cannot hold almost surely. Indeed, in principle I can replace in (1) the r.v. $\Phi\left(\frac{X(t)}{\sqrt{t}}\right)$ with $1 - \Phi\left(\frac{X(t)}{\sqrt{t}}\right)$ without altering the argumentation. This is for example, the idea used in antithetic Monte Carlo simulation.

Table 1 illustrates the procedure that allows to map the process X into the log-return process. Given the last step of the procedure in Table 1, the function g is then recovered via (1).

If log-return have a Gaussian distribution with mean $\mu(0, t)$ and variance $\sigma^2(0, t)$, then $F(x; t) = \Phi\left(\frac{x - \mu(0,t)}{\sigma(0,t)}\right)$, and the function g turns out to be linear

$$g(X(t), t) = \frac{\sigma(0, t)}{\sqrt{t}} X(t) + \mu(0, t).$$

---

[4] There is a slight difference in notation between the Erio's note and the present paper because the random variable $g(X)$ in the note is assumed to be normal, whereas in this paper, X is assumed to be normal. However, this difference can be easily reconciled, and the two formulas can be made consistent.

Therefore the idea behind the above construction is that the more non-linear the function g is, the further the distribution of r deviates from normality. It is worth noting that assuming a standard Brownian motion for X is not strictly necessary. In the case that the exact distribution of r is used for X, this will always result in a linear function for g.

The construction of the function g requires the knowledge of the cumulative distribution function F. In this regard, the seminal work by Breeden and Litzenberger (1978) is particularly relevant. They demonstrate how to derive the risk-neutral distribution of the stock price at a given horizon T using option prices,[5] through the following formula

$$Pr\left(S_T \leq x\right) = F_S\left(x; T\right) = \frac{1}{P(0, T)} \left.\frac{\partial c\left(K, T\right)}{\partial K}\right|_{K=x}. \tag{2}$$

where $c\left(K, T\right)$ is the call option price with strike K and time to maturity T, and $P(0, T) = e^{-r_f(0,T)T}$ is the discount factor relative to the option maturity T and $r_f(0, T)$ is the continuously compounded spot rate term structure. Therefore, the desired log-return CDF is

$$F(x; T) = F_S\left(S_0 e^x; T\right).$$

The cumulative distribution function has been derived from option prices but is currently only available on specific dates t_i, $i = 1, \ldots, n$ when option prices are observable, rather than on arbitrary dates. For practical considerations, it is essential to extend this function to any time t, where $0 < t \leq t_n$. In practice, a straightforward and effective method involves employing simple linear interpolation across maturities.

Other methods, such as kriging, spline interpolation, or local polynomial regression, are also available for interpolating the value of a function at a specific point by calculating a weighted average of known function values in the neighborhood of the point. Future research may explore the adoption of these methods, as well as more advanced techniques, such as shape-constrained kernel regression, as proposed in Yagi et al. (2020).

Let $G(X, t)$ denotes the linear interpolation of the function g across the maturity dimensions. Note that interpolation for times shorter (or longer) than the shortest (longest) available maturity is not implemented.

The mapping g between the standard Brownian motion (SBM) and the log-return guarantees that the martingale condition is automatically fulfilled on the specific dates t_i, $i = 1, \ldots, n$ for which option prices are available, rather than on arbitrary dates. By replacing $g(X, t)$ with the interpolating function $G(X, t)$, a function can be obtained that allows for the generation of the distribution of returns at any date, but the martingale restriction is not guaranteed at interpolated maturities. For this reason, the

---

[5] A completely ignored paper by Gilat (1971) had already illustrated back in 1971 the fact that the distribution of a r.v. Y is completely determined by the truncated moments $\mathbb{E}(Y - u)^+$ for u ranging over all real numbers.

log-return process needs to be redefined as

$$r(0, t) = m(0, t) + G(X(t), t), \tag{3}$$

instead of $r(0, t) = g(X(t), t)$ and $m : \mathcal{R}^+ \to \mathcal{R}$ is a deterministic function of time. The role of the deterministic function $m(0, t)$ is to guarantee that the usual martingale restriction holds, i.e.

$$\mathbb{E}_0(S(t)) = S(0)e^{(r_f(0,t) - q(0,t))t}, \ \forall t \geq 0,$$

where $q(0, t)$ is the dividend yield term structure (in general assumed flat at the level q). This condition can be guaranteed by imposing

$$m(0, t) = (r_f(0, t) - q(0, t))t - \ln \mathbb{E}_0\left(e^{G(X(t), t)}\right), \tag{4}$$

assuming that the expectation exists and is finite. Notice in particular, that if the function G is linear in X, due to the fact that the functions g are also linear and time independent, i.e. $g(X, t_i) = \sigma X, \ \forall i = 1, \ldots, n$, then, using the moment generating function of the normal distribution, I obtain the usual drift restriction

$$m(0, t) = \left(r_f(0, t) - q(0, t) - \frac{\sigma^2}{2}\right)t.$$

It is worth noting that since the function G is a convex combination of monotone increasing functions, it is itself monotone increasing. Therefore, it always ensures the recovery of a valid distribution function for the log-return. This fact rules out the existence of arbitrage opportunities across strikes. However, the investigation if the interpolating function G allows to generate option prices that exclude arbitrage opportunities across maturities is something that need to be better verified but beyond the scope of the present paper. To this regard the results in Carr and Madan (2005) are relevant. They show that the absence of call spread, butterfly spread and calendar spread arbitrages is sufficient to exclude all static arbitrages from a set of option price quotes across strikes and maturities on a single underlying. A check that can be conducted to offer initial assurance in the current construction is to ensure that the variance of the log-returns is a monotonically increasing function of time. If it were not the case, it would indicate the existence of a sub-period (t_i, t_{i+1}) in which the log-return $r(t_i, t_{i+1})$ has negative variance, which would clearly be unacceptable. Given that the variance of returns is related to the variance of the function G, I can compute for different horizons t

$$\mathbb{V}ar_0(G(X, t)) = \int_{-\infty}^{\infty} (G(x, t) - \mu(0, t))^2 \phi(x, 0, \sqrt{t}) dx,$$

where

$$\mu(0, t) = \int_{-\infty}^{\infty} G(x, t) \phi(x, 0, \sqrt{t}) dx,$$

and $\phi(x, 0, \sqrt{t})$ is the normal density function with zero mean and variance t. In this way, I can check that $\mathbb{V}ar_0\left(G(X, t)\right)$ is increasing in time. The calculation of the term structure of higher order moments, such as skewness and kurtosis, is also straightforward. Indeed, the moment of order j is

$$\mathbb{E}_0\left(r^j(t)\right) = \mathbb{E}\left(\ln^j(S_s/S_0)\right) = \int_{-\infty}^{+\infty} (m(0, t) + G(x, t))^j \phi(x, 0, \sqrt{t}) dx. \quad (5)$$

In the context of the risk-neutral specification, the function G is also helpful to define the payoff implicit in option prices. Instead of asking what is the implied risk-neutral distribution implicit in option prices, I can ask what is the payoff implicit in observed market option prices given that the Black–Scholes model holds. If the Black–Scholes model holds, I receive as payoff the amount

$$\left(S(0)e^{\sigma X(t) + (r_f(0,t) - q(0,t) - \sigma^2/2)t} - K\right)^+,$$

and the price I pay for it is given by the Black–Scholes formula. However, market option prices are in general not consistent with the Black–Scholes formula. The payoff that I have to receive in order to justify the observed option price, assuming the normality of returns, is

$$\left(S(0)e^{G(X(t),t) + m(0,t)} - K\right)^+$$

Therefore, the function G allows us to re-interpret deviation from Black–Scholes not in terms of a non-flat implied volatility surface as typically done, but by looking at the distance between implied and true payoff

$$\Delta(X(t), k, t) = \left(S(0)e^{G(X(t),t) + m(0,t)} - e^k\right)^+$$
$$- \left(S(0)e^{\sigma X + (r_f(0,t) - q(0,t) - \sigma^2/2)t} - e^k\right)^+.$$

The function $\Delta(X(t), k, t)$ is the increase/reduction in the payoff I receive at maturity t depending on the log-return and log-strike, given the observed option prices.

The manuscript by Erio also says that the Ito's lemma allows to obtain the (risk-neutral) stock price dynamics. Indeed, I have

$$dS(t) = (r_f(0, t) - q(0, t))S(t)dt + \sigma_S(S, t)S(t)dX(t),$$

and the local volatility function $\sigma_S(S, t)$ is given by

$$\sigma_S(S, t) = \left.\frac{\partial G(X, t)}{\partial X}\right|_{X = G^{-1}(\ln(S) - m(0,t), t)}. \quad (6)$$

By resorting to the derivative of the inverse function, I obtain

$$\frac{\partial G(X, t)}{\partial X} = \frac{\phi\left(\frac{X(t)}{\sqrt{t}}\right)}{f\left(\Phi\left(\frac{X(t)}{\sqrt{t}}\right); t\right)\sqrt{t}}.$$

Here, $\phi(z)$ is the standard normal probability density function, and $f(x, t)$ is the density probability function implied in option prices, i.e. the derivative of the return CDF $F(x, t)$. Therefore, the model is able to recover from option prices also the so-called local volatility function. It is interesting to compare the present approach to the Dupire formula, Dupire (1994), that is widely used in quantitative finance for its ability to recover the local volatility function from quoted put and call options, as follows:

$$\sigma_S(S, t) = \sqrt{\left.\frac{\frac{\partial C(K,t)}{\partial t} + (r_f - q)K\frac{\partial C(K,t)}{\partial K} + \frac{1}{2}2K^2\frac{\partial^2 C(K,t)}{\partial K^2}}{\frac{1}{2}K^2\frac{\partial^2 C(K,t)}{\partial K^2}}\right|_{K=S}}. \tag{7}$$

An advantage of the approach in the present paper is that the local volatility can be extracted using solely the first derivative of the function G with respect to x, whereas the Dupire formula necessitates the computation of the first partial derivative with respect to time and the first and second derivative with respect to the strike of the option price surface. Therefore, the Dupire formula requires an interpolation (and sometimes extrapolation) method for the option surface (or the implied volatility surface) to extend it to a continuous set of strikes and maturities and then the computation of the required derivatives. Thereafter, the risk-neutral price process can be simulated using the computed local volatility function. In general, the simulation needs some care, because the resulting transition law is not known and therefore the Euler scheme with a very small time step must be used. In our approach, this issue does not occur: the log-return is Markovian in the SBM X, so I can simulate the SBM in an exact way and then I recover the stock return via the mapping G. In practice, this approach fully avoids the need to estimate the local volatility function in (7).

However, the two approaches are not contradictory. In fact, the Dupire formula (7) can be used in conjunction with the framework presented here. Indeed, the model presented here can generate an option surface via the calculation of option prices using

$$c(K, t) = \mathbb{E}_0\left(\left(S_0 e^{m(0,t)+G(z\sqrt{t},t)} - K\right)^+\right)P(0, t), \tag{8}$$

where z is a standard normal r.v.. Then the local volatility function can be computed applying to (8) the Dupire formula. It is clear that, from a practical point of view, the application of (6) appears more direct with respect to the application of (7): only the derivative wrt X of the function G is required.

Finally, I can also rewrite the dynamics of the log-return as the solution of an autonomous stochastic differential equation. Given that the function G is continuous

and monotone increasing, it admits an inverse function, so that

$$X(t) = G^{-1}\left(r(0, t) - m(0, t), t\right).\tag{9}$$

Similarly, at time $t + \Delta t$,

$$X(t + \Delta t) = G^{-1}\left(r(0, t + \Delta t) - m(0, t + \Delta t), t + \Delta t\right).\tag{10}$$

Given that $X(t + \Delta t) = X(t) + z\sqrt{\Delta t}$, where $z \sim \mathcal{N}(0, 1)$, I have the following exact dynamics for the stock return

$$r(0, t + \Delta t) = m(0, t + \Delta t) + G\left(G^{-1}\left(r(0, t) - m(0, t), t\right) + z\sqrt{\Delta t}, t + \Delta t\right),$$

that establishes a relationship between $r(0, t)$ and $r(0, t + \Delta t)$. Therefore, even if the driving process X has independent and identically distributed (iid) increments, the return $r(t, t + \Delta)$ covering the period $[t, t + \Delta t]$ depends in a non linear way on the return up to time t, i.e. $r(0, t)$. Indeed, given the additivity property of log-returns, i.e. $r(0, t + \Delta t) = r(0, t) + r(t, t + \Delta t)$, I have

$$\begin{aligned} r(t, t + \Delta t) = {} & m(0, t + \Delta t) - r(0, t) \\ & + G\left(G^{-1}\left(r(0, t) - m(0, t), t\right) + z\sqrt{\Delta t}, t + \Delta t\right). \end{aligned}$$

Therefore, in general, the return process will show some path-dependence, i.e. the return $r(t, t + \Delta)$ has some memory because it depends in a non-linear way on the return up to time t, i.e. $r(0, t)$, as well as on the Gaussian increment $z\sqrt{\Delta t}$. An evident exception is the case where the function G is linear: in this case returns over not-overlapping periods are iid.

As final remark it is worth mentioning the relationship of the present approach with Goldenberg (1991). This author illustrates an invariance principle for pricing options on diffusion processes related by time and scale changes. He transforms non-Gaussian asset price data in order to make them Gaussian. He also presents a proposition showing that reducibility generates analytical option pricing formulas for complex underlying diffusion processes. The approach in Goldenberg (1991) is further investigated in Carr and Madan (2001) with the aim of fitting the volatility surface. The approach adopted in the present paper is different, because no preliminary assumption is made on the dynamics of the return process but it is simply recovered by imposing the link G with the SBM.

3 Recovering the distribution function at available maturities

This section illustrates a possible procedure for implementing the methodology outlined in the previous section. Given that it is a crucial input for constructing the function g, a method for extracting the distribution of returns from option prices is at first presented.

The process of constructing a risk-neutral density from option prices requires satisfying fundamental constraints, such as ensuring that the density is non-negative, integrates to one, and accurately prices all available calls and puts. However, constructing such a density requires option prices for a broad range of strikes and maturities, which may not be feasible in practical situations. Consequently, an infinite number of density functions could correspond to a given set of option prices over a finite range of strikes.

To address the challenges stemming from the discreteness of observed option prices, I utilize the Stochastic Volatility Inspired (SVI) parameterization of the implied volatility smile, as proposed by Gatheral and Jacquier (2014). The SVI parameterization is intended to capture the essential characteristics of the implied volatility smile, such as its curvature, skewness, and kurtosis. It is a versatile model that can be employed to fit the implied volatility smile for a broad range of underlying assets and maturities. One advantage of the SVI parameterization is that it is relatively simple and computationally efficient, making it a popular choice for practitioners in the financial industry. This approach involves a mathematical formula that characterizes the shape of the implied volatility curve for a given underlying asset as a function of log-moneyness $\kappa = \ln(K/S(0))$. For a given maturity T, the total variance $TV(k, T)$ relative to the log-moneyness k is given by

$$TV(\kappa, T) = a + b \left(\rho(\kappa - m) + \sqrt{(\kappa - m)^2 + \sigma^2} \right) \qquad (11)$$

The implied volatility $IV(k, T)$ is related to the total variance via

$$IV(\kappa, T) = \sqrt{\frac{IV^2(\kappa, T)}{T}}.$$

The SVI parametrization represents the implied volatility as a function of moneyness and time to expiration, using five parameters. The five parameters are:

- The at-the-money variance level a ($a \in \mathbb{R}$): increasing this parameter implies a vertical translation of the smile;
- The slope b of the wings, ($b \in \mathbb{R}^+$): increasing b increases the slope of both the put and call wings, tightening the smile.
- The curvature m of the wings ($m \in \mathbb{R}$);
- The skewness ρ of the smile ($|\rho| \in (0, 1)$): increasing ρ decreases (increases) the slope of the left (right) wing, a counter-clockwise rotation of the smile;
- The horizontal shift σ of the smile ($\sigma \in \mathbb{R}^+$): increasing σ reduces the ATM curvature of the smile. The case $\sigma = 0$ corresponds to a linear smile.

Let $c(K, T)$ denotes the Black–Scholes formula as function of the moneyness κ and the total variance $TV(\kappa, T)$

$$c(\kappa, T) = (0) \left(e^{-qT} N(d_1) - e^{\kappa} P(0, T) N(d_2) \right) \qquad (12)$$

where

$$d_{1,2} = \frac{\ln\left(\frac{e^{-qT}}{P(0,T)}\right) - \kappa \pm \frac{1}{2}TV(\kappa,T)}{\sqrt{TV(\kappa,T)}}.$$

and where the remaining parameters are: stock price S_0, the dividend yield q and the discount factor $P(0,T)$. I proceed as follows.

1. For a given maturity T having N_T option quotes, with $N_T \geq 10$, the SVI parametrization can be fitted to the observed volatility smile by solving a non-linear least square problem, i.e.

$$\min_{a,b,m,\rho} \sum_{i=1}^{N_T} (IV(\kappa_i, T) - IV_{mkt}(\kappa_i, T))^2. \tag{13}$$

2. The interpolated volatilities are converted back into the appropriate call and put prices by using the Black–Scholes formula (12).

3. Using the result from Breeden and Litzenberger (1978), the stock price CDF computed at $s = K$ is

$$F_S(K;T) = P(0,T)\frac{\partial c\left(\ln\left(\frac{K}{S_0}\right), T\right)}{\partial K} \tag{14}$$

$$= 1 - \mathcal{N}(d_2) + P(0,T)Kn(d_2)\sqrt{T}\frac{\partial TV\left(\ln\left(\frac{K}{S}\right), T\right)}{\partial K}\Bigg|_{K=s}, \tag{15}$$

where

$$\frac{\partial TV\left(\ln\left(\frac{K}{S}\right), T\right)}{\partial K} = \frac{b}{K}\left(\rho + \frac{\kappa - m}{\sqrt{(\sigma^2 + (\kappa - m)^2)}}\right)_{|\kappa = \ln\left(\frac{K}{S}\right)}$$

4. The CDF $F(x;T)$ of return is related to the stock price CDF $F_S(S_0e^x;T)$ via

$$F(x;T) = F_S(S_0e^x;T). \tag{16}$$

5. The desired monotonic transformation g can be recovered via

$$g(X,T) = F^{-1}\left(\Phi\left(\frac{X}{\sqrt{T}}\right);T\right),$$

where the inverse of the CDF is computed by using a look-up table with 4,000 equally spaced points in the interval $[-6, 6]\sqrt{T}$.

6. Steps 1–5 are repeated for all liquid maturities $T = t_1, \ldots, t_n$ having at least 10 options satisfying the constraints mentioned in the next section.

7. The interpolation procedure is applied to generate the surface $G(X, t)$ and to obtain the drift adjustment $m(0, t)$.
8. As additional output of the proposed procedure I can recover the local volatility function by (6) that can be computed via finite differences or by using (2) that requires the probability density function PDF $f_S(s; T)$ that can be computed by taking the derivative of the CDF F_S in (15).

4 An empirical illustration

In this section, I provide a step-by-step demonstration of the practical implementation of the proposed model.

4.1 Dataset

I considered American options for Apple Inc. on the closing date of April 20th, 2023, with maturities ranging from 9 to 331 days.[6] To ensure data quality, I filtered the total of 1034 options by only considering OTM calls and puts with a minimum volume of 4, a minimum open interest of 1, and a minimum price of 0.8. After applying these filters, I was left with a dataset of 201 options, consisting of 91 calls and 110 puts. The key characteristics of the dataset are summarized in Table 2, which also reports the risk-free rate derived from the term structure of SOFR rates. The underlying stock price was 167.62, and the dividend yield was 0.54%.[7]

4.2 Calibrating the SVI model

Using available option quotes, a binomial tree with 250 steps per year was employed, based on the Cox-Ross parametrization and the Snell envelope, to extract the implied volatility. Subsequently, the extracted implied volatilities were used to calibrate the SVI model. To guarantee a minimal statistical validity to the calibration procedure, maturities with less than 10 available options were excluded, resulting in a final set of maturities consisting of 23, 93, 121, 149, 184, 212, 240 and 275 days. It should be noted that longer maturities typically have a larger set of available strikes, which facilitates a better calibration of the SVI parametric model. For example, the 11 available strikes for options expiring in 93 days cover the range of 135 to 190, whereas for options expiring in 275 days, the range widens to 95–235.

The results of the calibration of the SVI model are presented in Table 3, where the mean square root error (MSRE) is reported in the last column. The fitting accuracy was generally high, as evidenced by the good agreement between the extracted implied volatilities and those obtained via the SVI model, with the exception of the 275-day maturity due to the presence of an outlier. Positive values of b indicated the non-flat

---

[6] Option prices were recovered on April 20th, 2023 via a python web scraper code from the Yahoo Finance web page https://finance.yahoo.com/quote/AAPL/options?fr=sycsrp_catchall.

[7] These quantities were recovered from the Yahoo Finance web page containing the summary information about Apple Inc. at https://finance.yahoo.com/quote/AAPL?p=AAPL.

Table 2 Dataset characteristics

Maturity (days)	Number of options	OTM calls	OTM puts	Smallest strike	Largest strike	Risk-free rate (%)
9	3	1	2	165	170	5.019
16	8	3	5	157.5	175	5.019
23	10	4	6	155	177.5	5.019
30	5	2	3	155	175	5.019
37	7	3	4	150	180	5.019
44	7	3	4	150	180	5.019
58	9	4	5	145	185	5.019
93	11	4	7	135	190	5.067
121	15	7	8	130	200	5.101
149	18	8	10	120	205	5.097
184	18	8	10	115	205	5.072
212	19	10	9	105	220	4.964
240	24	11	13	100	225	4.875
275	29	14	15	95	235	4.795
Sum	201	91	110			

Table 3 Calibrated parameters of the SVI parametrization across maturities

Days	a	b	m	ρ	σ	Rate	RMSE
23	−0.00212	0.02303	0.17177	0.00263	0.09208	0.05019	0.00521
93	−0.01369	0.07124	0.17454	0.01387	0.19221	0.05067	0.00158
121	−0.02169	0.09135	0.21146	0.01384	0.23742	0.05101	0.00108
149	−0.02841	0.08411	0.17632	−0.32778	0.35751	0.05097	0.00109
184	−0.04072	0.09488	0.16536	−0.42422	0.47398	0.05072	0.00072
212	−0.02879	0.11237	0.21982	−0.23674	0.40243	0.04964	0.00077
240	−0.03298	0.12122	0.21167	−0.28145	0.45053	0.04875	0.00071
275	−0.04322	0.13462	0.13128	−0.40119	0.55390	0.04795	0.00770

nature of the implied volatility curve for all maturities. Moreover, non-zero values of ρ, which are positive for short maturities and negative for long maturities, imply that the slope of the volatility curve may vary across different maturities becoming more flat for long horizons, see Fig. 1.

4.3 Construction of the functions $g(x, t_i)$ and $G(x, t)$

Following calibration of the SVI model, the implied distribution function for the liquid maturities was extracted using formulae (15)–(16). Subsequently, formula (1) was applied to obtain the functions $g(x, t_i)$, which are depicted in Fig. 2.

The curvature of the functions is highly significant, indicating a departure from the normal benchmark. The steepness of the curves increases for longer maturities, reflecting a wider range of potential returns. The benchmark, which is only shown for the longest maturity of 275 days, is a straight line. The benchmark in practice allows to capture exactly the mean and standard deviation of the implied distribution of log-returns. The function $G(X, t)$ is then obtained through linear interpolation over time to maturity and is presented in Fig. 3.

Imposing the martingale restriction

Next, the function m is determined using formula (4), and it is displayed in Fig. 4. Additionally, Fig. 5 shows the standard deviation (volatility) of cumulative returns (blue curve). For sake of comparison, I also provide the standard deviation assuming that returns follow a process with independent and identically distributed increments (red curve), which follows the square root rule and scales with $\sqrt{t}$. Since the recovered volatility function remains above the one consistent with the assumption of i.i.d. returns, it indicates that the recovered risk-neutral process exhibits some positive autocorrelation.

In Table 4 I perform two additional checks, by comparing the mean and the standard deviation of the function G with the ones obtained by numerical integration of the implied density. This check is performed for the different maturities and the agreement is very good if the range of the r.v. X is taken to be $[−6, 6] \sqrt{T}$. If the considered range is narrower, for instance, $[−4, 4] \sqrt{T}$, the accuracy is good but not entirely satisfactory.
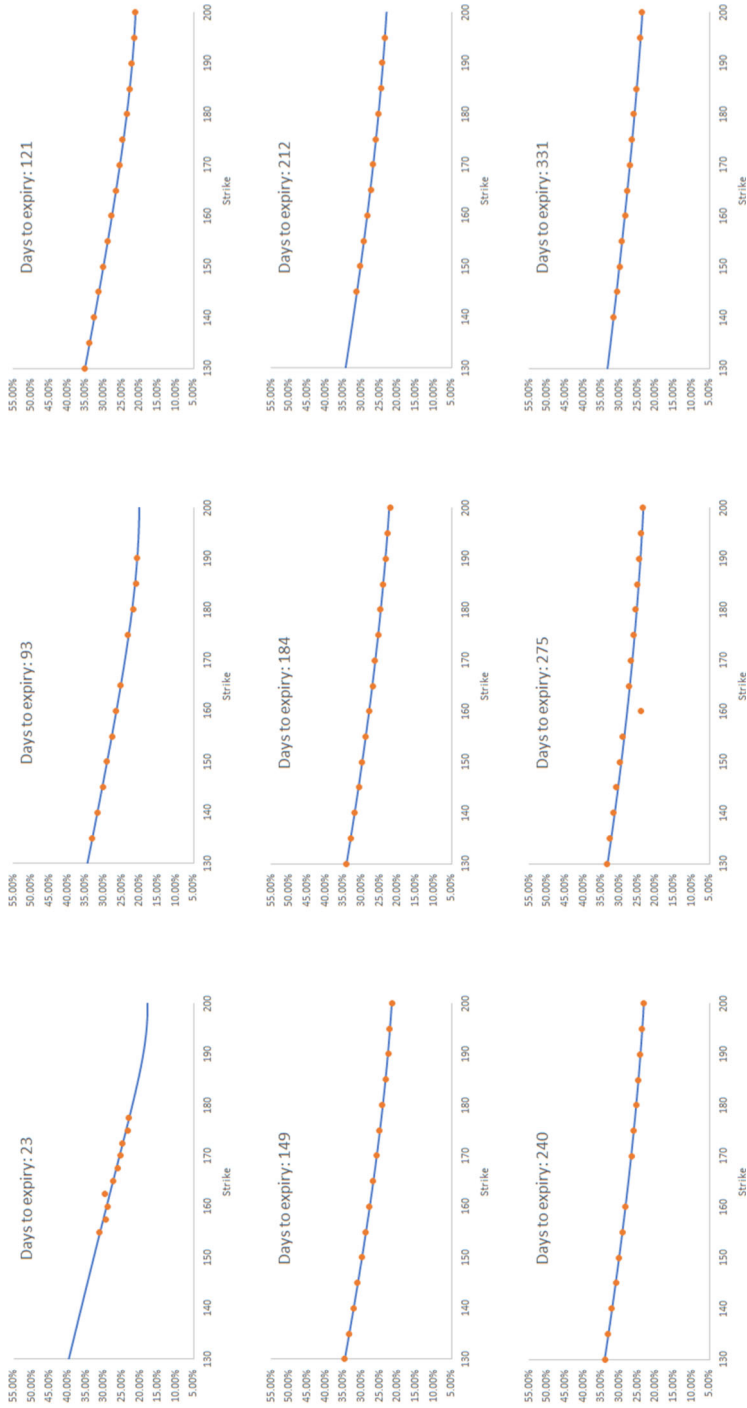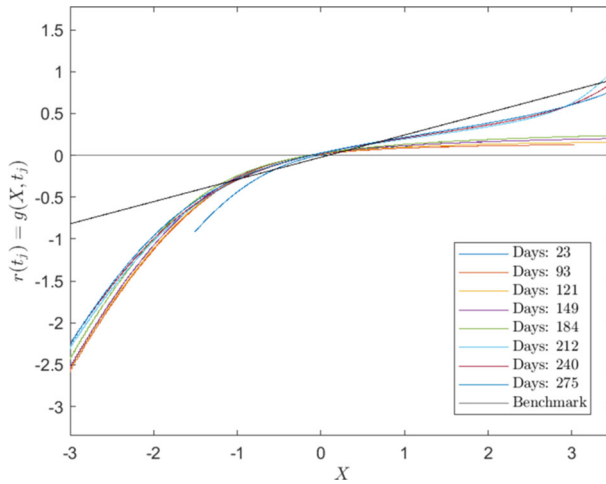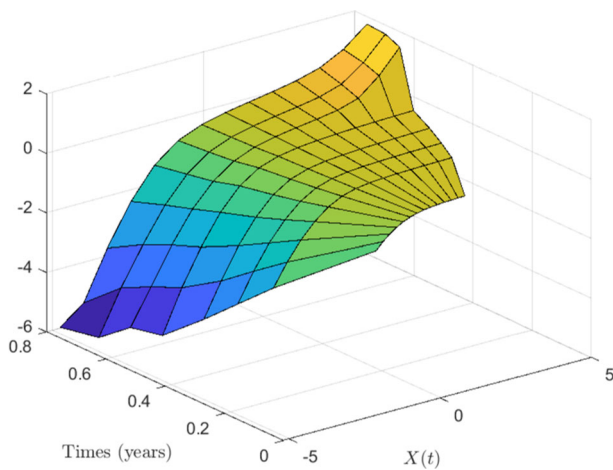
Fig. 1 Calibrated SVI parametrizations across different maturities. Orange dots: observed implied volatilities, Blue line: calibrated SVI curve

Fig. 2 Functions $g(X, t_i)$ for the considered maturities and varying X

Fig. 3 Interpolating function $G(X, t)$ for different times (in years) and values of X

Local volatility function

Fig. 6 illustrates the local volatility function $\sigma_S(S, t)$, computed using formula (6) across various option maturities. The local volatility function is plotted within the quoted strike range specified in Table 2 for each maturity. Notably, the Figure confirms the characteristic inverse relationship between volatility and stock price. The inverse relationship seems more pronounced for shorter option maturities in general.

Simulating the stock price
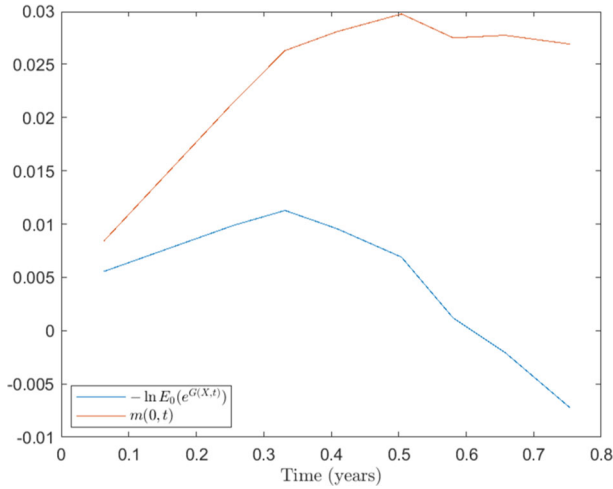
Fig. 4 Function $m(0, t)$
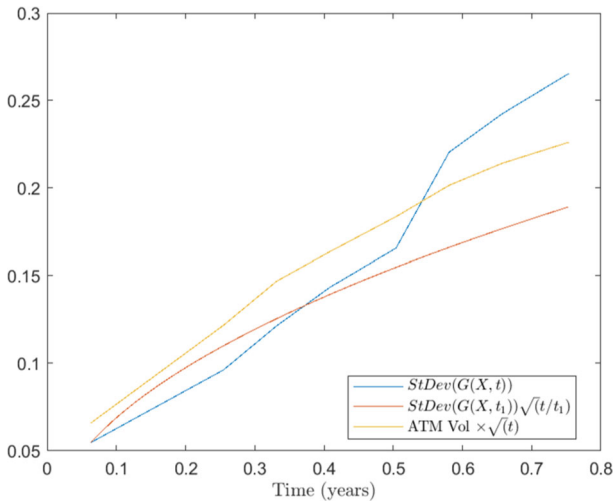
Fig. 5 Volatility of cumulative log-returns for different horizons and comparison with the square-root rule arising assuming log-returns are iid and the at-the-money implied volatility (scaled by the squared root of the time to maturity)

Once I have calibrated the model, it becomes straightforward to simulate the price process, by applying

$$S(t) = S(0)e^{G(X(t),t)+m(0,t)} = S(0)e^{(r-q)t} \frac{e^{G(X(t),t)}}{\mathbb{E}_0\left(e^{G(X(t),t)}\right)},$$

Table 4 Comparing the standard deviation of the implied risk-neutral distribution and the reconstructed G transformation for different option expires (in days)

Days	$\mathbb{S}Dev_0(r(0, T))$ (%)	$\mathbb{S}Dev_0(G(X(T), T)]$ $X \in [-4,\ 4]\sqrt{T}$ (%)	$X \in [-6,\ 6]\sqrt{T}$ (%)
23	5.47	5.47	5.47
93	9.63	9.60	9.63
121	12.15	12.11	12.14
149	14.34	14.29	14.34
184	16.58	16.52	16.58
212	22.05	22.00	22.05
240	24.25	24.19	24.25
275	26.55	26.47	26.55
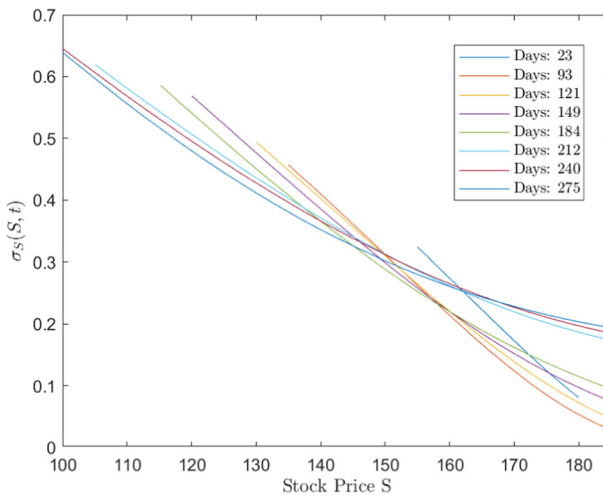
Fig. 6 Local volatility function

where the process $X(t)$ is simulated according to

$$X(t + \Delta t) = X(t) + z(t)\sqrt{\Delta},$$

with $z(t) \sim \mathcal{N}(0, 1)$ and starting value $X(0) = 0$.

Figure 7 displays 10,000 simulated stock price paths over a period of 22–275 days, with a daily time step. The simulated paths begin at the shortest available maturity (22 days) and end at the longest available maturity (275 days).

5 Final remarks

A natural question that arises is whether assuming standard Brownian motion dynamics for X is the only possible choice. However, it is clear by now that this is not the case.
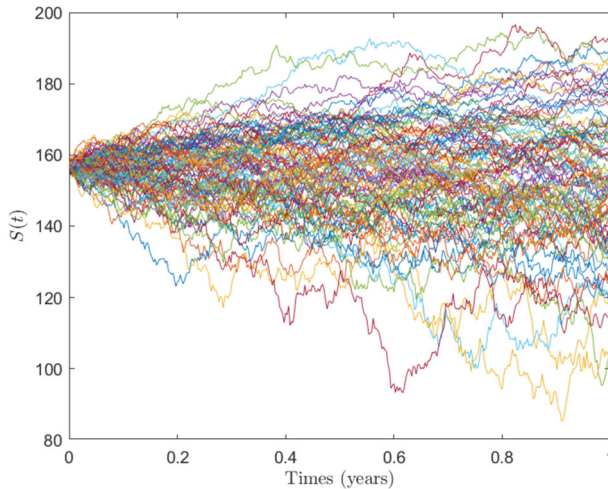
Fig. 7 Simulated trajectories of the stock price consistent with the observed market option prices

If a flat smile is observed in the market, the implied distribution is precisely normal, and the function g is linear. The more pronounced the smile, the more non-linear g becomes. Nevertheless, even if the smile is not flat, g can still be linear if I correctly guess the correct dynamics of X that generated the observed smile, such as a one-dimensional jump-diffusion process. In this case, the adopted distribution coincides with F, and g becomes linear. In other words, the extent to which g deviates from a linear function serves as a signal of how far the assumed model is from being accurate. However, adopting non-Gaussian dynamics for the latent factor would require fixing free parameters that appear in the dynamics of the more general model. Alternatively, I can assign arbitrary but sensible values to these parameters and obtain an acceptable fit to the volatility smile. I could then use the methodology presented here to exactly recover market option prices. In this case, the function g should not deviate too much from being linear.

A further possible investigation would be to replace the SBM continuous time process for X by a discrete time one. This extensions would represent an alternative approach to the implied tree approach in Rubinstein (1994) and in Derman and Kani (1998) and could represent a valid tool to deal with the early exercise feature.

Finally, several other applications that deserve attention for future research include pricing exotic derivatives, the potential for a latent variable with stochastic volatility to act as the driving factor, and extending the model to capture the dependence between multiple assets.

In conclusion, I hope that this short paper has been able to capture, at least in part, the intuition presented in Erio's note.

Declarations

Conflict of interest As the nature of this research did not involve any sensitive or potentially harmful topics, no ethical issues were encountered during the study, and therefore no ethical approvals or special permissions were required. Additionally, there were no potential conflicts of interest among the authors or funding sources. The research did not involve Human Participants and/or Animals.

References

Amerio, E., Fusai, G., Vulcano, A.: Implied volatility derivatives and risk neutral model for market volatility. Financ. Opt. Res. Centre—U. Warwick 126 (2003)

Breeden, D.T., Litzenberger, R.H.: Prices of state-contingent claims implicit in option prices. J. Bus. **51**(4), 621–651 (1978)

Carr, P., Madan, D.: Determining volatility surfaces and option values from an implied volatility smile. Quant. Anal. Financ. Mark, World Sci 163–191 (2001)

Carr, P., Madan, D.B.: A note on sufficient conditions for no arbitrage. Financ. Res. Lett. **2**(3), 125–130 (2005)

Castagnoli, E.: Associative operations and g-transformation. manuscript (2001)

Derman, E., Kani, I.: Stochastic implied trees: arbitrage pricing with stochastic term and strike structure of volatility. Int. J. Theor. Appl. Financ. **01**(01), 61–110 (1998)

Dupire, B.: Pricing with a smile. Risk **7**(1), 18–24 (1994)

Fusai, G.: Un'osservazione sulla compatibilità tra principio media varianza ed utilità attesa. In: Proceedings XVI Amases Conference, Treviso (1992)

Gatheral, J., Jacquier, A.: Arbitrage-free SVI volatility surfaces. Quant. Financ. **14**(1), 59–71 (2014)

Gilat, D.: Some conditions under which two random variables are equal almost surely and a simple proof of a theorem of Chung and Fuchs. Ann. Math. Stat. **42**(5), 1647–1655 (1971). Accessed 01 Mar 2024

Goldenberg, D.H.: A unified method for pricing options on diffusion processes. J. Financ. Econ. **29**(1), 3–34 (1991)

Modesti, P.: Dominanza stocastica e trasformazioni monotone. Studi matematici, Università Commerciale Luigi Bocconi. Ist Metodi Quant **15** (1993)

Rubinstein, M.: Implied binomial trees. J. Finance **49**(3), 771–818 (1994)

Yagi, D., Chen, Y., Johnson, A.L., Kuosmanen, T.: Shape-constrained kernel-weighted least squares: estimating production functions for Chilean manufacturing industries. J. Bus. Econ. Stat. **38**(1), 43–54 (2020)