Review

From Constant to Rough: A Survey of Continuous Volatility Modeling

Giulia Di Nunno [1,2] **, Kęstutis Kubilius** [3] **, Yuliya Mishura** [4,5] **and Anton Yurchenko-Tytarenko** [1,*]

1   Department of Mathematics, University of Oslo, 0851 Oslo, Norway; giulian@math.uio.no
2   Department of Business and Management Science, NHH Norwegian School of Economics, 5045 Bergen, Norway
3   Faculty of Mathematics and Informatics, Vilnius University, LT-03225 Vilnius, Lithuania; kestutis.kubilius@mif.vu.lt
4   Department of Probability, Statistics and Actuarial Mathematics, Taras Shevchenko National University of Kyiv, 01601 Kyiv, Ukraine; yuliyamishura@knu.ua
5   Division of Mathematics and Physics, Mälardalen University, 722 20 Västerås, Sweden
*   Correspondence: antony@math.uio.no

Abstract: In this paper, we present a comprehensive survey of continuous stochastic volatility models, discussing their historical development and the key stylized facts that have driven the field. Special attention is dedicated to fractional and rough methods: without advocating for either roughness or long memory, we outline the motivation behind them and characterize some landmark models. In addition, we briefly touch on the problem of VIX modeling and recent advances in the SPX-VIX joint calibration puzzle.

1. Introduction

Ultimately, finance evolves around the interplay between the expected return and risk. The established benchmark for the latter is *volatility*: loosely defined, this term refers to the degree of variability of an asset S in time which is traditionally quantified as

$$\sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(\log\frac{S(t)}{S(t-1)}\right)^2}. \tag{1}$$

The rationale behind using the *log*-returns in Equation (1) is simple and can be traced back to P. Samuelson: the prices are strictly positive and, hence, it is reasonable to represent them as an exponential of some random variable. For example, if one models the returns as

$$\frac{S(t)}{S(t-1)} = e^{\xi_t}, \quad t = 1, \dots, T,$$

where $\xi_t \sim \mathcal{N}(0, \sigma^2)$, $t = 1, \dots, T$, are centered independent and identically distributed random variables, the value of Equation (1) becomes a strongly consistent estimator of the *standard deviation* parameter σ.

However, it turns out that the metric Equation (1) is not as innocuous as it may seem at first sight. For example, if one computes Equation (1) using samples from distinct time periods, the results can be *drastically* different. In some sense, this is not a big surprise: there are periods on the market when prices exhibit minimal variation and other times when they

fluctuate with extreme intensity (especially during crises). Moreover, as noted in, e.g., [1,2], the variability in asset prices cannot be fully explained only by changes in "fundamental" economic factors. The important conclusion from this observation is as follows: **volatility changes over time in an unpredictable manner.**

In this survey, we offer an overview of one approach aimed at addressing this phenomenon: *stochastic volatility modeling*. This methodology, which can be traced back to the early discrete-time model of Clark (1973) [3], has evolved into an enormous, deep, and highly engaging area of research. Contrary to the older surveys of Skiadopoulous (2001) [4], Shepard and Andersen (2009) [5], and Duong and Swanson (2011) [6], we put a significant emphasis on fairly recent topics of fractional and rough volatility. We tried to make the presentation as friendly as possible, with the hope that readers familiar with stochastic analysis will find most parts of the material accessible.

Before proceeding to the main part of the paper, let us make two important disclaimers.

- First of all, apart from more straightforward applications in, e.g., mean-variance portfolio theory originating from [7] by H. Markowitz, the concept of volatility has a crucial role in another significant area within finance, i.e., *non-arbitrage pricing theory*, conceptualized by F. Black, M. Scholes, and R. Merton in [8,9]. This theory has a unique and intricate perspective on volatility and has been the major driving force in volatility modeling for the past four decades. Therefore, in this survey, we consider the volatility predominantly from the option pricing viewpoint.

- Our focus primarily centers on *continuous stochastic volatility models in continuous time*. This means that we will not discuss various discrete-time models such as ARCH, GARCH, or EARCH of Engle [10], Bollerslev [11], and Nelson [12], or multiple jump-diffusion models including the approaches of Duffie et al. [13] or Barndorff-Nielsen and Shephard [14]. The omission of jumps seems to be the most notable gap in light of the realism this approach provides. However, given our desire to highlight fractional and rough volatility in more detail, the inclusion of discrete-time and jump-diffusion models with the level of detail they deserve would substantially inflate the size of this survey and result in a notable change in emphasis. Therefore, we leave these two topics for a separate work. Readers with a specific interest in discrete-time models are referred to [15,16]; for jump models, see the specialized books by Barndorff-Nielsen and Shepard [17], Gatheral [18], Rachev et al. [19], or Tankov and Cont [20]. For other modeling viewpoints that are different from stochastic volatility, see the overviews given by Shiryaev [21] and Mariani and Florescu [22].

The paper is structured as follows. In Section 2, we present a historical context for stochastic volatility modeling as well as listing several relevant stylized facts that are aimed to be reproduced. We devote separate attention to the notion of *implied volatility*, since the latter serves as the focal point of the field. Section 3 gives a detailed survey of continuous stochastic volatility models from the constant elasticity of variance (CEV) model proposed by J. C. Cox in 1975 to the modern fractional and rough volatility approaches. For each class of models, we provide the context of their applicability and characterize their advantages and shortcomings. In Section 4, we touch on selected aspects of a special topic in volatility modeling: VIX. After a brief detour in its computation, we describe two problems within VIX modeling: reproducing positive VIX skew and the SPX-VIX joint calibration problem. In addition, we analyze advances in continuous stochastic volatility modeling for solving these problems. Section 5 concludes our presentation.

2. Market Volatility: Empirical Challenges and Stylized Facts

2.1. From Random Walk to Geometric Brownian Motion

Probability theory and stochastic analysis have evolved into irreplaceable tools for economics and finance. The first attempts to model financial markets using probabilistic methods can be traced back to the French economist Jules Regnault (for more details about this lesser known and somewhat underestimated person and his contributions to finance, the reader is referred to [23]): as early as 1863 [24], he proposed modeling stock prices with,

what we call nowadays, symmetric random walks. However, the foundation for employing probability in finance is commonly attributed to another French mathematician, Louis Bachelier, and his 1900 dissertation titled "*Théorie de la spéculation*" [25], where he considered market modeling and even derivative pricing in a very comprehensive manner. Naturally, the tools of stochastic calculus had not yet been formulated at that time, which means that the dissertation might lack some modern aspects of mathematical rigor. Nonetheless, if one translates Bachelier's reasoning into today's mathematical language (an excellent job in contextualizing Bachelier's work from the perspective of modern mathematics was achieved by M. Davis and A. Etheridge in [26]), it turns out that Bachelier essentially modeled stock price dynamics with a prototype of a standard Brownian motion 5 years before A. Einstein and his famous paper [27].

Unfortunately, Bachelier's ideas did not gain immediate recognition and went relatively unnoticed. Only more than 50 years after the publication of the original dissertation, statistician Jimmy Savage inexplicably stumbled upon this work and brought it to the attention of a number of researchers in economics. One of those researchers turned out to be Paul Samuelson, who described Savage's finding in the foreword to the English translation of Bachelier's dissertation [26] as follows:

> "...*Discovery or rediscovery of Louis Bachelier's 1900 Sorbonne thesis* [...] *initially involved a dozen or so postcards sent out from Yale by the late Jimmie Savage, a pioneer in bringing back into fashion statistical use of Bayesian probabilities. In paraphrase, the postcard's message said, approximately, 'Do any of you economist guys know about a 1914 French book on the theory of speculation by some French professor named Bachelier?'*
>
> *Apparently I was the only fish to respond to Savage's cast. The good MIT mathematical library did not possess Savage's 1914 reference. But it did have something better, namely Bachelier's original thesis itself.*
>
> *I rapidly spread the news of the Bachelier gem among early finance theorists. And when our MIT PhD Paul Cootner edited his collection of worthy finance papers, on my suggestion he included an English version of Bachelier's 1900 French text...*"

Samuelson noticed that Bachelier's quasi-Brownian model could potentially take negative values, which is an unrealistic property for real-life prices. Therefore, he proposed (Samuelson himself acknowledged—see, e.g., his comments in a foreword to [26]—that the same idea was independently expressed by an astronomer M. Osborne in [28]) a simple but very useful modification of Bachelier's original approach: Brownian dynamics should be used to model *price logarithms* rather than the *prices themselves*. After a small adjustment with a linear trend, Samuelson's model took the form of a *geometric* or (*economic relative*, the term used by Samuelson himself [29]) *Brownian motion*

$$S(t) = S(0) \exp\left\{ \left(\mu - \frac{\sigma^2}{2}\right)t + \sigma W(t) \right\}, \quad \mu \in \mathbb{R}, \quad S(0), \, \sigma > 0, \tag{2}$$

or, as a stochastic differential equation (SDE),

$$dS(t) = \mu S(t)dt + \sigma S(t)dW(t). \tag{3}$$

Note that the term *volatility* obtains a very specific meaning within the model Equations (2) and (3); namely, it refers to the parameter σ. Such terminology is very intuitive and, moreover, agrees well with the metric Equation (1) mentioned above: if $0 = t_0 < t_1 < ... < t_n = T$ is a partition of the interval $[0, T]$, then

$$\frac{1}{T} \sum_{k=1}^{n} \left(\log \frac{S(t_{k+1})}{S(t_k)} \right)^2 \to \sigma^2$$

in L^2 as the diameter of the partition $\max_k |t_{k+1} - t_k| \to 0$ (see, e.g., [30]). Due to its simplicity and tractability, the log-normal process Equations (2) and (3), together with the

corresponding notion of volatility, subsequently became a mainstream choice for stock price models for the next couple of decades. Even now, well informed of multiple arguments against the geometric Brownian motion, practitioners still use it as a benchmark or a reliable "first approximation" model.

It is important to note that, in addition to the market model, Bachelier also considered the problem of option pricing and eventually derived an expression that can be called a precursor of the now famous *Black–Scholes formula*. Of course, his rationale was not based on the no-arbitrage principle and had a number of shortcomings, as is often the case with pioneering works. The correction of those shortcomings became the subject of a number of studies in the 1960s, among which one can mention [31–33]. Samuelson himself also heavily contributed to that topic; see, e.g., [34] or his paper [35] (in co-authorship with Robert Merton), where it was suggested to consider a warrant/option payoff as a function of the price of the underlying asset. One could argue that these works were just a few steps away from the breakthrough made by Black, Scholes, and Merton just a couple of years later.

Here, it is worth paying attention to the fact that the options market remained relatively illiquid until the end of the 1960s. The reason for that was the lack of a consistent pricing methodology, and serious investors regarded options as akin to gambling rather than worthy trading instruments. It is somewhat ironic that even Robert Merton himself, right in his pivotal article [9], wrote the following:

"Because options are specialized and relatively unimportant financial securities, the amount of time and space devoted to the development of a pricing theory might be questioned..."

However, in 1968, a demand for that type of contract suddenly arose from the Chicago Board of Trade. The organization observed a significant decline in commodity futures trading on its exchange and, therefore, opted to create additional instruments for investors. They settled on options, and, in 1973, the Chicago Board of Options Exchange commenced its activity. Precisely in that year, two revolutionary papers appeared: *"The pricing of options and corporate liabilities"* [8] by Fischer Black and Myron Scholes and *"Theory of rational option pricing"* [9] by Robert Merton. As a side note, the publication of the Black and Scholes paper was far from a smooth process: in 1987 [36], Black recalled that the manuscript was rejected first by the *Journal of Political Economy* and then by the *Review of Economics and Statistics*. The paper was published only after Eugene Fama and Merton Miller personally recommended the *Journal of Political Economy* to reconsider its decision (in the meanwhile, Robert Merton showed a great deal of academic integrity by delaying the publication of his own article so that Black and Scholes would be the first).

The ideas of Black, Scholes, and Merton revolutionized mathematical finance and enjoyed empirical success: Stephen Ross, for instance, claimed in 1987 [37] that

"When judged by its ability to explain the empirical data, option pricing theory is the most successful theory not only in finance, but in all of economics."

2.2. Implied Volatility and the Smile Phenomenon

The main result of Black, Scholes, and Merton can be formulated as follows: if a stock follows the model Equations (2) and (3), then, under some assumptions, the discounted *no-arbitrage price* of a standard European call option $C^{\text{B-S}} = C^{\text{B-S}}(t, S)$ evolves as a function of the current time t and current price S and must satisfy a partial differential equation, known now as the *Black–Scholes formula*, of the form

$$\frac{\partial C^{\text{B-S}}}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C^{\text{B-S}}}{\partial S^2} + rS\frac{\partial C^{\text{B-S}}}{\partial S} - rC^{\text{B-S}} = 0 \tag{4}$$

with a boundary condition (in what follows, we will utilize the standard notation $(x)_+ := \max\{x, 0\}$)

$$C^{\text{B-S}}(T, S) = (S - K)_+, \tag{5}$$

where r denotes the instantaneous interest rate, that is assumed to be constant, T is the maturity date of the option, and K is its exercise price. As mentioned above, the Black–Scholes–Merton rationale relies on a number of rather abstract assumptions that do not align with real-world market conditions. In particular, their reasoning required specific price dynamics, the absence of transaction costs, and the capacity to buy and sell any quantity of assets. However, the inability to perfectly replicate the reality does not necessarily carry significant implications. Black, Scholes, and Merton themselves were aware of the limitations of their approach: for example, ref. [38] quotes Fisher Black on this subject:

"Yet that weakness is also its greatest strength. People like the model because they can easily understand its assumptions. The model is often good as a first approximation, and if you can see the holes in the assumptions you can use the model in more sophisticated ways."

What really mattered was the successful empirical performance of the vanilla Black–Scholes–Merton model; as noted by J. Wiggins, one of the pioneers of continuous-time stochastic volatility modeling, *"given the elegance and tractability of the Black–Scholes formula, profitable application of alternate models requires that economically significant valuation improvements can be obtained empirically"* [39].

However, after the Black Monday market crash in 1987, it became evident that there were glaring flaws within the log-normal paradigm, prompting the need for rectification. Jackwerth and Rubinstein [40] described the problem as follows:

"Following the standard paradigm, assume that stock market returns are lognormally distributed with an annualized volatility of 20% (near their historical realization). On 19 October 1987, the two month S&P 500 futures price fell 29%. Under the lognormal hypothesis, this is a -27 standard deviation event with probability 10^{-160}. Even if one were to have lived through the entire 20 billion year life of the universe and experienced this 20 billion times (20 billion big bangs), that such a decline could have happened even once in this period is a virtual impossibility."

Evidently, experiencing *"virtually impossible"* price falls which rendered investors insolvent was already a good argument to reassess financial modeling approaches. Yet, apart from that single shock that, in principle, could be attributed to a single anomaly, there was another consistent phenomenon that manifested itself in the aftermath of Black Monday: **the volatility smile**.

In order to explain the problem, let us first discuss the parameters of the Black–Scholes model in more detail. Note that the sole value in the Black–Scholes formula, Equations (4) and (5), that is not directly observable is the volatility parameter σ: indeed, the maturity date T and exercise price K are given in the specifications of the given option contract, whereas the price $S(t)$ can be read from the market. The volatility σ is unknown and necessitates some form of estimation based on data. In 1976, Latané and Rendelman [41] proposed an elegant method to address this problem. First of all, note that Equations (4) and (5) have an explicit solution of the form

$$
\begin{aligned}
C^{\text{B-S}}(t, S(t)) = {} & S(t)\Phi\left(\frac{\log\frac{e^{r(T-t)}S(t)}{K} + \frac{\sigma^2}{2}(T-t)}{\sigma\sqrt{T-t}}\right) \\
& - Ke^{-r(T-t)}\Phi\left(\frac{\log\frac{e^{r(T-t)}S(t)}{K} - \frac{\sigma^2}{2}(T-t)}{\sigma\sqrt{T-t}}\right),
\end{aligned}
\tag{6}
$$

where $\Phi(x) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{y^2}{2}} dy$. For simplicity, let us take $t = 0$ and re-parametrize Equation (6) as

$$C^{\text{B-S}}(T, S(0), K, \sigma) := S(0)\Phi\left(\frac{\log \frac{e^{rT}S(0)}{K} + \frac{\sigma^2}{2}T}{\sigma\sqrt{T}}\right) - Ke^{-rT}\Phi\left(\frac{\log \frac{e^{rT}S(0)}{K} - \frac{\sigma^2}{2}T}{\sigma\sqrt{T}}\right). \quad (7)$$

Take the *actual* market price $C(T, K)$ of the corresponding option with payoff K and maturity date T and notice that, since $C^{\text{B-S}}(T, S(0), K, \sigma)$ is supposed to coincide with C, the volatility σ can be obtained from the equation

$$C^{\text{B-S}}(T, S(0), K, \sigma) - C(T, K) = 0. \quad (8)$$

The solution $\hat{\sigma} = \hat{\sigma}(T, K)$ to this equation is called the *implied* volatility. Note that in the literature (see, e.g., [42], p. 244), $\hat{\sigma}$ is often written as a function $\hat{\sigma}_{\text{log-m}}$ of time to maturity T and *log-moneyness* $\kappa := \log \frac{K}{e^{rT}S(0)}$, i.e.,

$$\hat{\sigma}_{\text{log-m}}(T, \kappa) := \hat{\sigma}(T, S(0)e^{\kappa+rT}).$$

In what follows, we will slightly abuse the notation and omit the subscript "log-m", using $\hat{\sigma}(T, K)$ for the function of strike and $\hat{\sigma}(T, \kappa)$ for the function of log-moneyness.

If the stock price model, Equations (2) and (3), indeed corresponds to reality well enough, $\hat{\sigma}(T, K)$ should be approximately constant for options with the same underlying asset but differing maturities T and strikes K. However, this is not the case! For instance, $\hat{\sigma}(T, K)$ turns out to change with T for fixed K (see Figure 1 below).

SPX implied volatility as function of maturity, K=4100
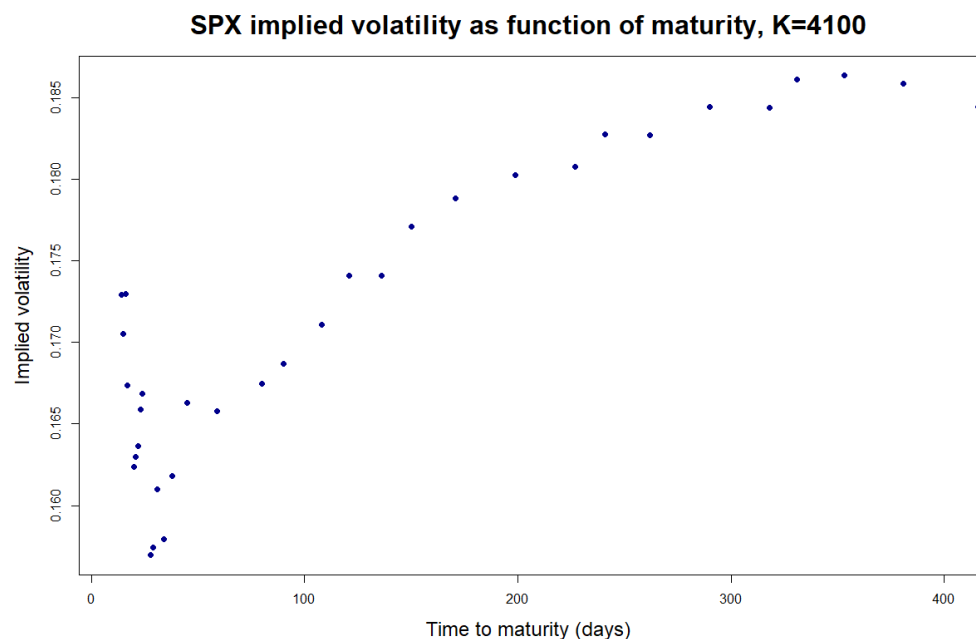
Figure 1. Variation in implied volatility with time to maturity T for a fixed strike K. The values were calculated using S&P500 (SPX) option prices on 3 May 2023, retrieved from *finance.yahoo.com* [43].

A straightforward modification of Equation (2), considered by Merton in his original paper [9], appears to accommodate this issue. Namely, if the volatility $\sigma = \sigma(t)$ is assumed to be a deterministic function of time, one can obtain a version of Equation (6) of the form

$$
\begin{aligned}
C^{\text{B-S}}(t, S(t)) = {} & S(t)\Phi\left(\frac{\log\frac{e^{r(T-t)}S(t)}{K} + \frac{1}{2}\int_t^T \sigma^2(s)ds}{\sqrt{\int_t^T \sigma^2(s)ds}}\right) \\
& - Ke^{-r(T-t)}\Phi\left(\frac{\log\frac{e^{r(T-t)}S(t)}{K} - \frac{1}{2}\int_t^T \sigma^2(s)ds}{\sqrt{\int_t^T \sigma^2(s)ds}}\right) \\
:= {} & S(t)\Phi\left(\frac{\log\frac{e^{r(T-t)}S(t)}{K} + \frac{1}{2}\overline{\sigma}^2(t,T)T}{\overline{\sigma}(t,T)\sqrt{T}}\right) \\
& - Ke^{-r(T-t)}\Phi\left(\frac{\log\frac{e^{r(T-t)}S(t)}{K} - \frac{1}{2}\overline{\sigma}^2(t,T)T}{\overline{\sigma}(t,T)\sqrt{T}}\right),
\end{aligned}
$$

where $\overline{\sigma}^2(t,T) := \frac{1}{T}\int_t^{t+T} \sigma^2(s)ds$. Then, the counterpart (at $t = 0$) of Equation (8) takes the form
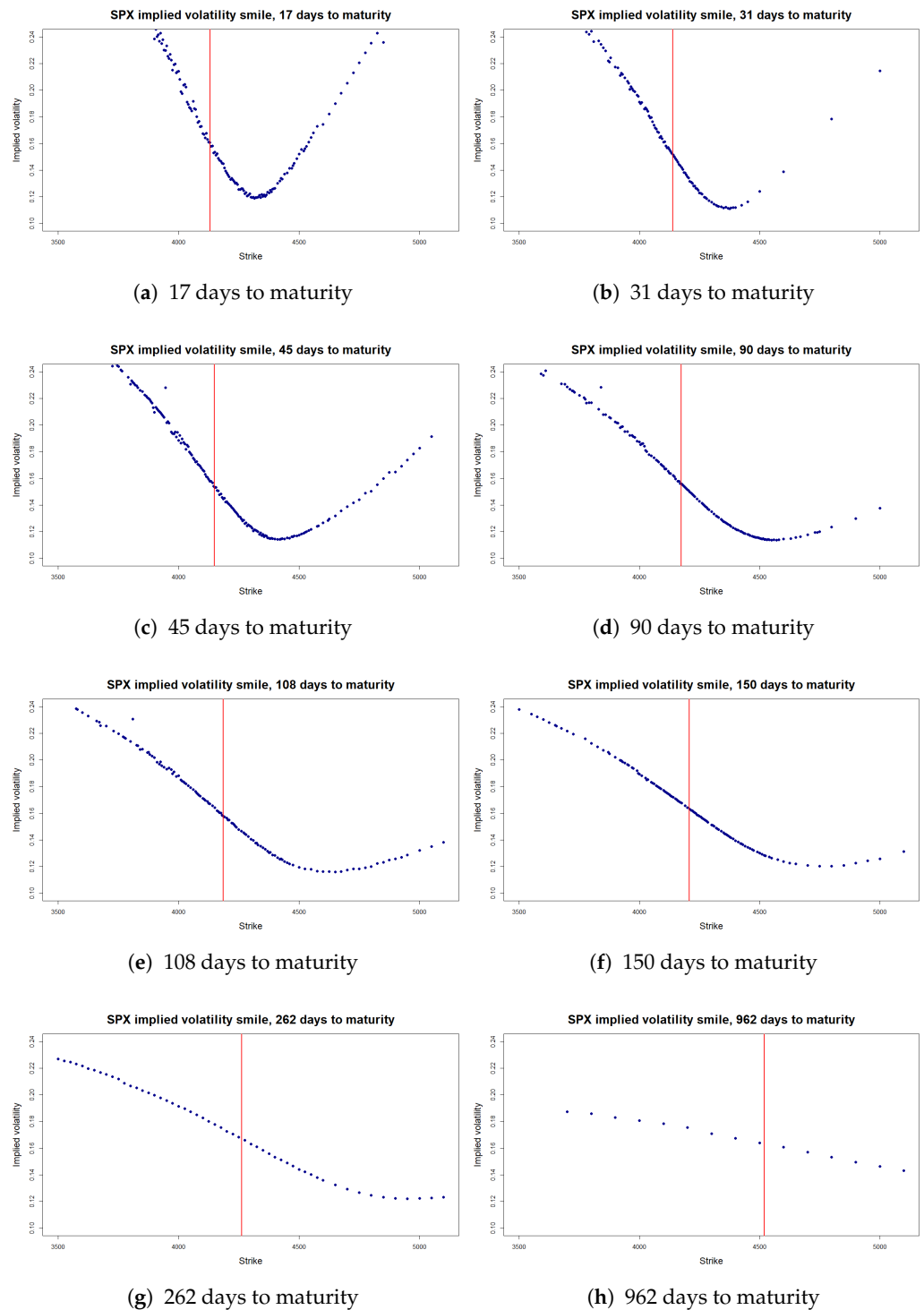
$$
C^{\text{B-S}}(T, S(0), K, \overline{\sigma}(0,T)) - C(T, K) = 0,
$$

so its solution $\widehat{\sigma}(T, K)$ is supposed to represent $\overline{\sigma}(0, T)$ and, hence, it is allowed to vary with T for fixed K. One may even argue that it is actually *reasonable* to assume that σ changes with time: as noted in [44], p. 144, "*there is nothing inconsistent about expecting high volatility this year and low volatility next year*".

Regarding the variation in K for a fixed T, the implied volatility remained relatively flat (more precisely, some dependence was present, but it was subtle enough to be disregarded; see, e.g., the discussion in [44], chapter 1) before the above-mentioned Black Monday crash in 1987. Starting from that (terrible) date, investors observed notable variability in the implied volatility in K characterized by distinct convex patterns (see Figure 2 below as well as [45]) which were eventually called "*volatility smiles*" or "*volatility smirks*". Such behavior was consistent, had a direct adverse impact on the empirical performance of the Black–Scholes formula, and could not be explained by the price dynamics of Equations (2) and (3).

Figure 2 also reveals two important typical characteristics of implied volatility smiles related to the change in the smile shape with T.

- First of all, note that the smile gradually flattens out as the time to maturity T increases. However, this decrease is fairly slow: for example, in Figure 2, the curvature is noticeable 262 days (Figure 2g) and even 962 days (Figure 2h) before maturity. The latter effect turns out to require special treatment; in this regard, see Section 3.3 below, as well as [46–48].
- Second, observe the behavior of the smile *at-the-money*, i.e., when $K = S(0)e^{rT}$ (or, in terms of log-moneyness, when $\kappa = 0$): as T becomes smaller, the smile at-the-money tends to become *steeper*. Figure 3 characterizes this phenomenon in more detail: for each T, we took implied volatilities of seven options with strikes closest to $S(0)e^{rT}$, performed a least squares linear fit on them, and depicted the absolute values of the resulting slopes in Figure 3a. It turns out that the variation in absolute slopes with T for shorter maturities seems to be well described by the *power law* $CT^{-\frac{1}{2}+H}$ with $H \approx 0$, *at least as a first approximation* (see the discussion at the end of Section 3.3 below, as well as [49,50]). As an illustration, the red lines in Figure 3a,b depict power-law fits for the SPX implied volatility slopes corresponding to 3 May 2023; for our dataset, $H = 0.06226572$. A similar type of behavior is reported by, e.g., Fouque, Papanicolaou, Sircar, and Sølna (2004) [51], Gatheral, Jaisson, and Rosenbaum (2018) [52] and, more recently, by Delemotte, De Marco, and Segonne (2023) [50].

(a) 17 days to maturity

(b) 31 days to maturity

(c) 45 days to maturity

(d) 90 days to maturity

(e) 108 days to maturity

(f) 150 days to maturity

(g) 262 days to maturity

(h) 962 days to maturity

Figure 2. Implied volatility smiles for different times to maturity. The values were calculated using S&P500 (SPX) option prices on 3 May 2023, retrieved from *finance.yahoo.com* [43]. Vertical red lines correspond to the at-the-money level $K = S(0)e^{rT}$.

(**a**) Regular scale   (**b**) Logarithmic scale
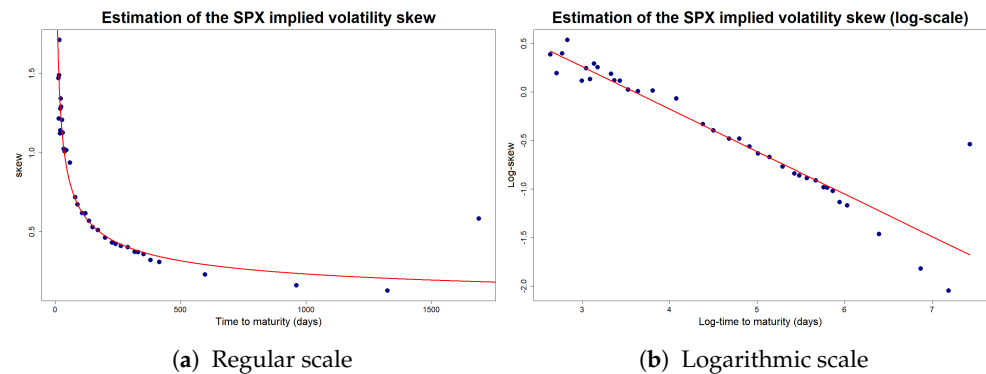
Figure 3. Absolute values of at-the-money implied volatility slopes on regular (**a**) and log–log (**b**) scales. Blue dots denote the data and red lines denote the regression fits $e^{1.5753587} T^{-0.4377343}$ (**a**) and $1.5753587 - 0.4377343 \log T$ (**b**). The values were calculated using S&P500 (SPX) option prices on 3 May 2023, retrieved from *finance.yahoo.com* [43].

In principle, a "*perfect*" model should capture the shape of the entire *implied volatility surface* $(T, K) \mapsto \widehat{\sigma}(T, K)$ (or, equivalently, $(T, \kappa) \mapsto \widehat{\sigma}(T, \kappa)$; see Figure 4). Here, by "*capturing the volatility surface*", we mean the established benchmark for assessing the performance of a model by its ability to reproduce the shape of $\widehat{\sigma}(T, K)$ or, equivalently, $\widehat{\sigma}(T, \kappa)$. As a rule, testing adheres to the following algorithm:

(1) the parameters of the model are calibrated with respect to the market data;
(2) the calibrated parameters are plugged into the model and the respective model no-arbitrage European option prices are computed;
(3) these "synthetic" prices are then used to construct the implied volatility surface by the procedure described here in Section 2.2;
(4) finally, the obtained model-generated surface is compared to the implied volatility surface extracted from real-world option prices.
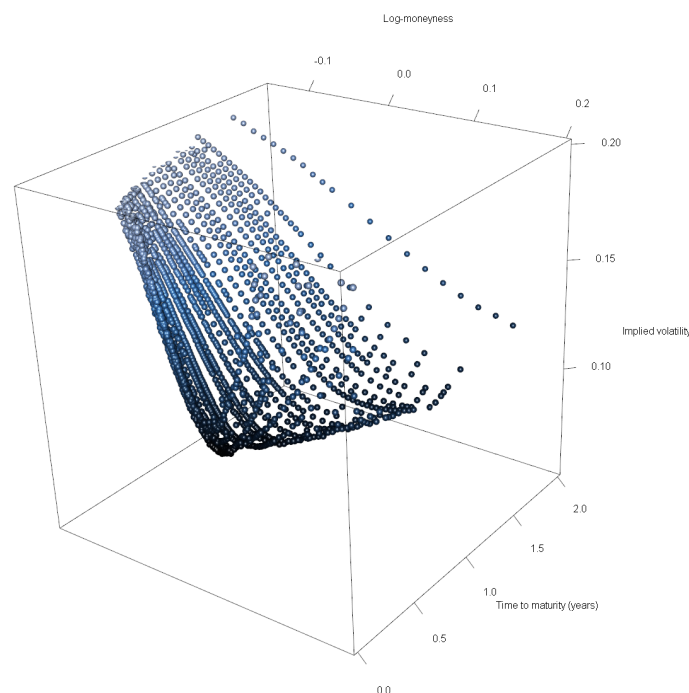
Figure 4. Implied volatility surface $(T, \kappa) \mapsto \widehat{\sigma}(T, \kappa)$. The values were calculated using S&P500 (SPX) option prices on 3 May 2023, retrieved from *finance.yahoo.com* [43].

For example, if one wants to reproduce the power law described above, one may demand the model-generated *at-the-money volatility skew*

$$\Psi(T) := \left|\frac{\partial}{\partial\kappa}\widehat{\sigma}(T,\kappa)\right|_{\kappa=0} \tag{9}$$

to have the power-law asymptotics $O(T^{-\frac{1}{2}+H})$, $T \to 0$. However, as we will see later, the task of constructing a model that reproduces all of the stylized facts simultaneously—both for long and short maturities—is not straightforward at all.

2.3. Fat Tails, Leverage, Clustering, and Long Memory

The empirical evidence of volatility smiles makes a spectacular point against the geometric Brownian dynamics of Equations (2) and (3). However, it certainly does not stand alone as the sole argument in this regard. In fact, objections to the log-normality of prices appeared as early as the log-normal model itself. In this section, we briefly enumerate several pivotal stylized facts concerning log-returns and the volatility of financial time series. For a more detailed discussion of this topic, we also refer our readers to [53] Section 2.2, [54] Section 3, well-known survey articles [1,55], or the book [56].

2.3.1. Fat Tails and Non-Gaussian Distribution of log-Returns

First of all, statistical analyses of price returns pointed out that their distributions have fat tails, i.e., the probabilities of extreme values tend to be significantly higher than predicted by log-normal models. In this context, we mention the empirical studies of Mandelbrot [57] (1963) and Fama [58] (1965); see also [59] for an early review. In response to this phenomenon, Mandelbrot proposed modeling price log-returns with α-stable distributions, which are characterized by infinite variance. However, this idea seems to contradict subsequent studies (such as [55]) which suggest that the variance of returns should be finite. Overall, ref. [55] gives the following summary regarding the properties of the "*true*" log-returns distribution: it tends to be non-Gaussian, sharp-peaked, and heavy-tailed. Clearly, there are multiple parametric models satisfying these three properties and one can mention log-return models based on inverse Gaussian distributions [60], generalized hyperbolic distributions [61], truncated stable distributions [62,63], and so on. For a more recent analysis of the topic, see also [64].

2.3.2. Leverage and Zumbach Effect

Another interesting phenomenon not grasped by the geometric Brownian motion is the so-called *leverage effect*: a negative correlation between the variance and returns of an asset. This empirical artifact, initially noticed by Black [65] and then studied in more detail by Christie [66], Cheung and Ng [67], and Duffee [68], was explained by Black himself as follows: a decrease in a stock price results in a drop of firm's equity value and hence increases its financial *leverage* (in this context, the term "leverage" means the company's debt relative to its equity). This, according to Black, makes the stock riskier and, hence, more volatile. Interestingly, the name "leverage effect" stuck due to this explanation, although subsequent research [69–71] pointed out that this correlation may not be connected to the leverage at all. Zumbach, in [56] section 3.9.1, gives a different interpretation of this negative correlation: downward moves of stock prices are generally perceived as unfavorable, triggering many sales and thereby increasing the volatility. Conversely, upward moves do not result in such drastic changes in investors' portfolios, given that the majority of market participants already hold long positions. Zumbach [72] also studied another effect of the same nature now recognized as the *Zumbach effect*: pronounced *trends* in stock price movement, irrespective of sign, increase the subsequent volatility. This effect stems from the fact that large price moves motivate investors to modify their portfolios, unlike scenarios where prices oscillate within a narrower range.

2.3.3. Volatility Clustering

The next empirical contradiction to the Black–Scholes–Merton framework arises from a direct econometric analysis of financial time series uncovering clusters of high and low volatility episodes. As noted by Mandelbrot [57], "*large changes tend to be followed by large changes, of either sign, and small changes tend to be followed by small changes*" (see Figure 5b).

There are several ways to interpret this phenomenon. Some authors (see, e.g., [73], chapter 3) identify volatility clustering with *mean reversion*: in loose terms, this term refers to volatility's tendency to "return back" to the mean level of its long-run distribution (see, e.g., [73] section 2.3.1 and chapter 3), with possible extended periods of staying above or below the latter. Another—and, perhaps, more nuanced—way to quantify clustering lies in analyzing the autocorrelation function of absolute log-returns (see, e.g., an excellent review [74] on the topic), i.e.,

$$\mathrm{corr}(|R(t)|, |R(t+\tau)|), \tag{10}$$

where the log-return $R(t) := \log\left(\frac{S(t+\Delta)}{S(t)}\right)$ is defined for some given time scale Δ (which may vary between a fraction of a second for tick data to several days). Multiple empirical studies [1,15,63,74–79] report that the autocorrelation function, Equation (10), is consistently positive and, moreover, shows signs of slow decay of the type $O(\tau^{-\beta})$, $\tau \to \infty$, with an exponent $0 < \beta \le 0.5$.

(**a**) Daily index data (**b**) Logarithmic returns
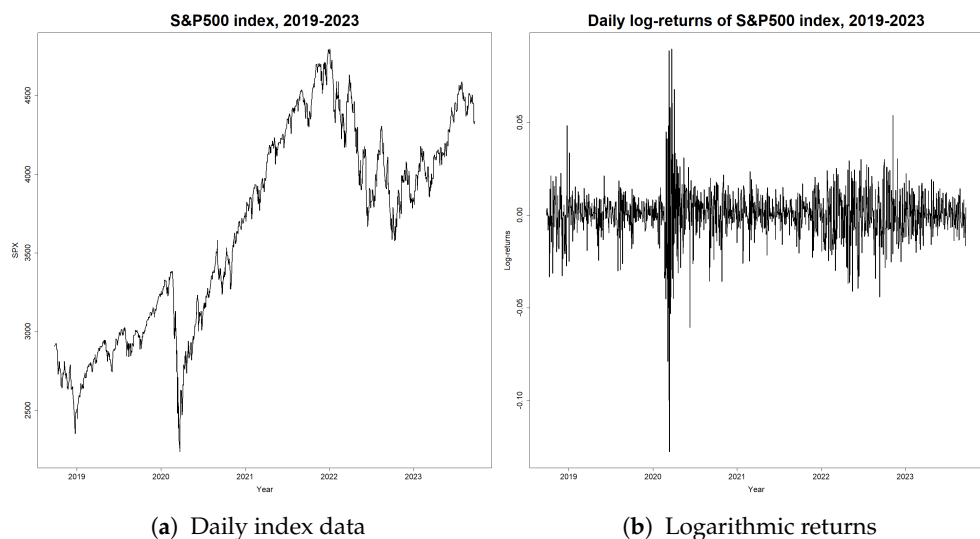
Figure 5. Daily split-adjusted values of the S&P500 (SPX) index (**a**) and the corresponding logarithmic returns (**b**). Note that the amplitude of fluctuations in log-returns tends to form clusters over time; the period of the highest variation, in March 2020, corresponds to the shock caused by the COVID-19 pandemic. The data were retrieved from *finance.yahoo.com* [43].

2.3.4. Long Range Dependence

The $O(\tau^{-\beta})$ decay of Equation (10) as $\tau \to \infty$ requires a separate discussion due to its profound implications. For $\beta \in (0, 1)$, such behavior is often referred to as the *long-range dependence* (see, e.g., [80,81]), and if its statistical significance is established, it indicates the *presence of memory* in the market. Notably, confirming the long-range dependence poses a substantial challenge: by its nature, memory manifests itself when $\tau \to \infty$, which raises the concern that any statistical estimation procedure of Equation (10) might be inconsistent due to, e.g., the non-stationarity of financial time series over extended time periods (see, e.g., discussion in [82] section 1.4). Moreover, as shown in [52] section 4, mis-specification of the model in the estimation procedure can result in a spurious long memory conclusion, even if the "*correct*" model does not exhibit long memory at all. Nevertheless, several studies still report the presence of long-range dependence on financial markets. For instance,

Willinger et al. [83] apply the so-called *rescaled range (R/S) analysis* technique to the CRSP (Center for Research in Security Prices) daily stock time series and find some weak evidence of memory in the data (as the authors write, "*...we find empirical evidence of long-range dependence in stock price returns, but because the corresponding degree of long-range dependence ... is typically very low ... the evidence is not absolutely conclusive*"). Lobato and Velasco [84] analyze volatility in connection to trading volumes and find that both of these financial characteristics exhibit the same degree of long memory. Another interesting point comes from the analysis of the implied volatility surface: Comte and Renault [47] noticed that the decrease in the smile amplitude as time to maturity increased was much slower than many advanced market models predicted. They argued that such an effect could be mimicked by having long memory in volatility (see also a simulation study [48] that directly confirms this claim).

3. Continuous Models of Volatility

As previously discussed, the real-world data does not support the standard log-normal framework of Equations (2) and (3). Luckily, developments in the option pricing theory subsequent to the seminal Black, Scholes, and Merton papers allowed for some flexibility in terms of the choice of price models. Namely, we refer to the gradual translation of the Black–Scholes–Merton approach into the language of martingale theory which evolved in the celebrated *Fundamental Theorem of Asset Pricing*—the result which connects non-arbitrage pricing and the existence of equivalent local martingale measures. In this regard, we mention the early research of Ross [85], Harrison and Kreps [86], Harrison and Pliska [87], Kreps [88], as well as subsequent seminal works of Delbaen and Schachermayer [89–91] (see also [92] for a detailed historical overview on the subject). This line of research eventually evolved into a general theory allowing for quite a broad variety of price models to choose from—hence giving researchers all the necessary tools to adjust the classical model, Equation (2), to account for volatility smiles and all other empirical inconsistencies.

As highlighted in Section 2, the particular problem of Equation (3) lies in the fact that the volatility parameter σ cannot be constant: it varies with time in an unpredictable manner, is correlated with the current price level, and has clusters of low and high values with some evidence of long-range dependence. One of the possible ways to treat this problem is by modifying Equation (3) by introducing an appropriate *stochastic volatility process* $\{\sigma(t),\ t \geq 0\}$ instead of the deterministic coefficient σ, i.e., by taking

$$dS(t) = \mu(t)S(t)dt + \sigma(t)S(t)dW(t). \tag{11}$$

Note that, in general, it is reasonable to treat the drift coefficient $\mu = \{\mu(t),\ t \geq 0\}$ as a stochastic process as well; however, we will not cover modeling drift in this survey.

Naturally, now the problem comes down to the selection of a particular process $\{\sigma(t),\ t \geq 0\}$ that, once inserted in Equation (11), can best reproduce the behavior of real-world prices. In this section, we list prominent continuous volatility modeling approaches and briefly characterize their performance.

3.1. Local Volatility

3.1.1. CEV Model

The first cluster of models covered in this survey are the so-called *local* or *deterministic volatility models*. The core idea behind this approach is the assumption that the process $\sigma = \{\sigma(t),\ t \geq 0\}$ in Equation (11) is a non-random function of the current price and time, i.e., $\sigma(t) = \sigma(t, S(t))$. Perhaps the first model of this kind was the *constant elasticity of variance (CEV)* model suggested by Cox in his 1975 note [93]. (Note that the original 1975 note remained unpublished since J.C Cox viewed it as a simple extension of his earlier work with S. Ross [94]. The version [93] cited here appeared in the *Journal of Portfolio Management*

in 1997 as a tribute to F. Black and his contribution to studying the leverage effect.) The CEV model assumes that $\sigma(t, S(t)) = \theta S^{\beta}(t)$, i.e., Equation (11) takes the form

$$dS(t) = \mu S(t)dt + \theta S^{1+\beta}(t)dW(t), \tag{12}$$

where the parameter β is called the *elasticity parameter*. Initially, Cox considered $\beta \in [-1, 0)$. In such a case, the increase in $S(t)$ decreases the value of $\sigma(t, S(t)) = \theta S^{\beta}(t)$ and vice versa, hence, reproducing the leverage effect. The case $\beta > 0$, more inherent to commodity prices, was discussed in [95]. For more details on the model Equation (12), we also recommend the survey article [96].

As a final remark regarding the CEV model, we note that the SDE Equation (12) exhibits substantially different behavior depending on the exact value of β. If $\beta \neq 0$, the diffusion coefficient in Equation (12) is not Lipschitz and, hence, the existence and uniqueness of the solution cannot be guaranteed by the classical result. For $\beta \in \left[-\frac{1}{2}, 0\right)$, the solution exists in the celebrated Yamada–Watanabe theorem (see, e.g., [97]), but the cases $\beta \in \left(-1, -\frac{1}{2}\right)$ and $\beta > 0$ require a separate and very careful treatment. For more details in that regard, we refer the reader to [98] chapter 5, as well as to the discussion in [99].

3.1.2. Local Volatility and Dupire Formula

The general local volatility was initially considered by Derman and Kani [100] and Dupire [101]. Both assumed the risk-neutral dynamics of the form

$$dS(t) = rS(t)dt + \sigma(t, S(t))S(t)dW(t), \tag{13}$$

where r denotes the interest rate. Note that this model is far more intricate and flexible than it seems at first glance. Unlike the CEV approach, Equation (13) does not specify any parametric form of the function σ. Instead, the no-arbitrage principle and properties of diffusion allow σ to be fully recovered from option prices using the celebrated Dupire formula (for its derivation, see the original paper [101] or [102] subsection 2.2.1), assuming no dividends,

$$\sigma(t, S) = \sqrt{2 \frac{\frac{\partial C}{\partial T}(T, K) + rS\frac{\partial C}{\partial K}(T, K)}{S^2 \frac{\partial^2 C}{\partial K^2}(T, K)}}\Bigg|_{(T,K)=(t,S)}, \tag{14}$$

where $C(T, K)$, as before, denotes the price at $t = 0$ of the European call-option with payoff K and maturity T. It is possible to prove (see, e.g., [102] subsection 2.2.2) that no-arbitrage conditions imply that the expression on the right-hand side of Equation (13) is well defined. Moreover, recalling that

$$C(T, K) = C^{\text{B-S}}(T, S(0), K, \widehat{\sigma}(T, K)), \tag{15}$$

where $C^{\text{B-S}}$ is defined by Equation (7), one can plug the right-hand side of Equations (15) to (14) in place of $C(T, K)$ and obtain the theoretical relation between the implied volatility $\widehat{\sigma}$ and local volatility function σ that can then be used to accurately calibrate Equation (13) to the current implied volatility smiles (see, e.g., [102] section 2.3). For more details on local volatility models, we refer the reader to [102] chapter 2.

3.1.3. Shortcomings of Local Volatility

As mentioned above, local volatility models are perhaps the simplest and the most straightforward generalizations of the original log-normal model Equations (2) and (3). In addition to their simplicity and tractability, they inherit another convenient property of the classical geometric Brownian motion: one can show that the market produced by Equation (13) is complete, i.e., any option with S as an underlying asset can be perfectly hedged by a self-financing portfolio composed exclusively of S and a riskless asset. However, despite the obvious mathematical attractiveness of this feature, some sources actually

view completeness as a disadvantage from the modeling viewpoint. As noted in [20] chapter 10,

> "*While absence of arbitrage is both a reasonable property to assume in a real market and a generic property of many stochastic models, market completeness is neither a financially realistic nor a theoretically robust property. From a financial point of view, market completeness implies that options are redundant assets and the very existence of the options market becomes a mystery, if not a paradox in such models.*"

This observation suggests that local volatility models may be too restrictive to grasp the complexity of the stock dynamics. And, importantly, such a claim is supported by multiple studies. For example, Buraschi and Jackwerth [103] utilize a formal statistical testing procedure to check whether options are indeed "*redundant*", and their findings strongly reject this hypothesis, questioning the viability of the model Equation (13). Another empirical study by Dumas et al. [104] analyzes the predictive performance of Equation (13) and concludes that its out-of-sample results are no better than ones of the standard Black–Scholes approach. Gatheral et al. [52] indicate that the dynamics Equation (13) tends to generate future volatility surfaces "*completely unlike those we observe*". Aït-Sahalia and Jacod [105] utilize a statistical test based on local time to check whether stock prices follow the SDE of the form

$$dS(t) = a(S(t))dt + b(S(t))dW(t),$$

where a and b are some deterministic functions, and report a "*clear rejection*" of such a model, advocating for alternative approaches (in fact, they consider an even more general framework, with jumps and microstructure noise, and still reject the local volatility hypothesis). For more details on the performance of local volatility models, see also an overview in section 4 of [4].

3.2. Stochastic Volatility Models

As indicated above, local volatility models can be subject to criticism as they might not possess sufficient flexibility to grasp the market in its full complexity. An alternative approach, known as *stochastic volatility*, involves modeling with a dedicated stochastic process that exhibits only partial dependence with S. Note that the term "stochastic volatility" may seem too vague given that the volatility $\sigma(t, S(t))$ employed in local models is also a stochastic process; nevertheless, in the literature, the term "stochastic" generally refers to approaches that treat volatility separately from the price dynamics. For example, $\sigma = \{\sigma(t), t \geq 0\}$ may be another diffusion driven by a separate Brownian motion $B = \{B(t), t \geq 0\}$ correlated with W. Naturally, within this framework, there are countless candidates for σ, and choosing a particular one that characterizes the market well is a complex task. However, stylized facts presented in Section 2.3 offer several starting points that are useful to keep in mind.

- To begin with, given the nature of volatility, it seems reasonable to model it with a non-negative process. In practice, this can be achieved by modeling log volatility and then taking an exponential. Alternatively, one may utilize non-negative diffusions such as Bessel-type processes (see, e.g., [30] chapter XI).
- Selecting an appropriate dependence structure between B and W from Equation (11) provides all the necessary tools to account for the leverage effect: a typical assumption is $\mathbb{E}[W(t)B(t)] = \rho t$ with $-1 < \rho < 0$.
- Another property shared by multiple stochastic volatility models is *mean-reversion*. As mentioned above in Section 2.3, this behavior seems to be a common trait for real-life volatility. Moreover, a thoroughly chosen mean-reverting process can, to some extent, mimic the clustering effect (see, e.g., [73], Chapter 3). A common approach to introduce mean reversion to the dynamics is to take a drift term of the form

$$\theta_1(\theta_2 - \sigma(t))$$

in the SDE for volatility. This drift *"pulls σ back"* to the level θ_2 whenever σ deviates from it. The parameter θ_1 calibrates the speed of mean-reversion.

Starting from the 1987 pioneering works of Hull and White [106], Wiggins [39], and Scott [107], multiple generations of approaches and numerous models have emerged in the literature. While not attempting to provide an exhaustive list, we present below a selection of notable contributions.

- Hull and White [106] assumed that the squared volatility $\sigma^2 = \{\sigma^2(t),\ t \geq 0\}$ is itself a geometric Brownian motion, i.e., price and volatility satisfy stochastic differential equations of the form

$$dS(t) = \mu S(t)dt + \sigma(t)S(t)dW(t),$$
$$d\sigma^2(t) = \theta_1 \sigma^2(t)dt + \theta_2 \sigma^2(t)dB(t)$$

  respectively, where B and W are two Brownian motions that are allowed to be correlated to account for the leverage effect. Note that the volatility process is positive but not mean-reverting.

- Wiggins [39] suggested a slightly more general dynamics of the form

$$dS(t) = \mu S(t)dt + \sigma(t)S(t)dW(t),$$
$$d\sigma(t) = f(\sigma(t))dt + \theta\sigma(t)dB(t).$$

- Scott [107] and Stein and Stein [108] considered the volatility to be an Ornstein–Uhlenbeck process, i.e.,

$$dS(t) = \mu S(t)dt + \sigma(t)S(t)dW(t),$$
$$d\sigma(t) = \theta_1(\theta_2 - \sigma(t))dt + \theta_3 dB(t). \tag{16}$$

  Note that σ is not positive: the Ornstein–Uhlenbeck process is Gaussian and, hence, can take negative values with positive probability. In practice, this issue is treated by either taking the absolute value of σ or introducing a reflecting barrier to the volatility dynamics [109].

- Heston [110] introduced the SDE of the form

$$dS(t) = \mu S(t)dt + \sqrt{\sigma(t)}S(t)dW(t),$$
$$d\sigma(t) = \theta_1(\theta_2 - \sigma(t))dt + \theta_3\sqrt{\sigma(t)}dB(t). \tag{17}$$

  In this model, now commonly referred to as the *Heston model*, the volatility follows the so-called *Cox–Ingersoll–Ross* or *square root* process (see also [111]), which enjoys strict positivity provided that $2\theta_1\theta_2 \geq \theta_3^2$. Benhamou et al. [112] considered a modification of Equation (17) of the form

$$dS(t) = \sqrt{\sigma(t)}S(t)dW(t),$$
$$d\sigma(t) = \theta_1(\theta_2(t) - \sigma(t))dt + \theta_3(t)\sqrt{\sigma(t)}dB(t),$$
$$d\langle W, B\rangle_t = \rho(t)dt,$$

  with time-dependent θ_2, θ_3, and correlation ρ to account for structural changes on the market. Another modification of Equation (17) was considered in [113]: there, the discounted price dynamics was assumed to follow

$$dS(t) = \sqrt{\sigma(t)}S(t)dW(t),$$

$$d\sigma(t) = \theta_1(Z_t)(\theta_2(Z_t) - \sigma(t))dt + \theta_3(Z_t)\sqrt{\sigma(t)}dB(t),$$

$$d\langle W, B\rangle_t = \rho(Z_t)dt,$$

where Z is a homogeneous continuous-time Markov chain that represents market switching between different regimes.

- Melino and Turnbull [114,115] considered the model of the form

$$dS(t) = (a + bS(t))dt + \sigma(t)S^\beta dW(t),$$
$$d\log(\sigma(t)) = (\theta_1 + \theta_2 \log(\sigma(t)))dt + \theta_3 dB(t), \tag{18}$$

where $\beta \in \left[\frac{1}{2}, 1\right]$, i.e., the log volatility is a mean-reverting Ornstein–Uhlenbeck process and S can be regarded as a combination of the CEV Equation (12) with stochastic volatility.

- Hagan et al. [116] proposed the stochastic *alpha-beta-rho (SABR) model* of the form

$$dS(t) = \sigma(t)S^\beta(t)dW(t),$$
$$d\sigma(t) = \alpha\sigma(t)dB(t). \tag{19}$$

where $\alpha \geq 0$, $0 \leq \beta \leq 1$ and $\mathbb{E}[W(t)B(t)] = \rho t$, $-1 < \rho < 1$. Note that the case $\beta \in \left(0, \frac{1}{2}\right)$ is quite intricate from the mathematical perspective and requires a special treatment; see [98] for more details.

- Lewis [117] and, later, Carr and Sun [118] (see also Baldeaux and Badran [119] section 2) considered the so-called *3/2 model*

$$dS(t) = S(t)\sqrt{\sigma(t)}dW(t),$$
$$d\sigma(t) = \kappa\sigma(t)(\theta - \sigma(t))dt + \varepsilon\sigma^{3/2}(t)dB(t). \tag{20}$$

The motivation behind the model partially comes from the statistical analysis of volatility models performed by Jawaheri [120] and Bakshi et al. [121], as well as from its ability to replicate the VIX skew (see Section 4 for further details). For more details on the existence and properties of the solution to Equation (20), we refer the reader to [98].

- Gatheral [122] presented the double CEV dynamics of the form

$$dS(t) = S(t)\sqrt{\sigma(t)}dW(t),$$
$$d\sigma(t) = -\kappa(\sigma(t) - \sigma'(t))dt + \eta\sigma^\alpha(t)dB(t), \tag{21}$$
$$d\sigma'(t) = -\kappa'(\sigma'(t) - \theta)dt + \eta'\sigma'^\beta(t)dB'(t),$$

where $\alpha, \beta \in \left[\frac{1}{2}, 1\right]$.

- Fouque et al. [123] propose a *multiscale* stochastic volatility. Their approach lies in modeling volatility with two diffusions, one fluctuating on a fast "time scale", and the other fluctuating on a "slow" time scale. Namely, their model takes the form

$$dS(t) = \mu S(t)dt + \sigma(t)S(t)dW(t),$$
$$\sigma(t) = f(Y(t), Z(t)),$$
$$dY(t) = \frac{1}{\varepsilon}(\theta_1 - Y(t))dt + \frac{\theta_2\sqrt{2}}{\sqrt{\varepsilon}}dB_1(t)$$
$$dZ(t) = \delta c(Z(t)) + \sqrt{\delta}g(Z(t))dB_2(t),$$

where f is a bounded positive function, $\delta, \varepsilon > 0$ are assumed to be small, Y is the fast scale volatility factor, i.e., a fast mean-reverting process, and Z is the slow scale volatility factor. Later, in [124], Fouque and Saporito employed the same multiscale paradigm to model the Heston-type stochastic volatility-of-volatility parameter:

$$
\begin{aligned}
dS(t) &= S(t)\sqrt{\sigma(t)}dW(t), \\
d\sigma(t) &= \kappa(\theta - \sigma(t))dt + \eta(t)\sqrt{\sigma(t)}dW'(t), \\
\eta(t) &= f(Y(t), Z(t)), \\
dY(t) &= \frac{\sigma(t)}{\varepsilon}\alpha(Y(t))dt + \sqrt{\frac{\sigma(t)}{\varepsilon}}\beta(Y(t))dB(t), \\
dZ(t) &= \sigma(t)\delta c(Z(t))dt + \sqrt{\delta\sigma(t)}g(Z(t))dB'(t).
\end{aligned}
\tag{22}
$$

In addition to reproducing the leverage effect and clustering via mean reversion, Brownian stochastic volatility models turn out to have an additional important advantage: they have an ability to capture, to some extent, "smiley" patterns of the implied volatility (see, e.g., [125], [73] section 2.8.2 or [18]). However, by design, parametric stochastic volatility models impose structural constraints on the relationship between the dynamics of the spot and implied volatilities and, hence, may not be able to capture the exact shape of the implied volatility surface (see, e.g., [18] figure 3.6).

A detailed analysis on that matter can be found in [42,126], in particular, as stated in [126] section 7.1 or [42] remark 11.3.21, Brownian diffusion models of volatility tend to produce the at-the-money skew $\Psi(T) = O(1)$, $T \to 0$, where Ψ is defined by Equation (9). This directly contradicts the empirical behavior of $O(T^{-\beta})$, $\beta \approx \frac{1}{2}$, mentioned in Section 2.3. In addition, as highlighted by Comte and Renault [47] (see also [46]), the decrease in the real-life volatility smile amplitude as $T \to \infty$ seems to be much slower than predicted by the classical Brownian diffusions. Comte and Renault connect this phenomenon to the *long-range dependence* in the volatility dynamics, which is well aligned with several empirical studies listed above in Section 2.3 that also report long memory on the market.

In other words, the inherent properties of Brownian motion limit the modeling capabilities of diffusion models and it is no surprise that a lot of effort was made to advance the stochastic volatility framework further to account for the mentioned inconsistencies.

3.3. Fractional and Rough Models

One of the ways to extend the stochastic volatility models described in Section 3.2 involves substituting the standard Brownian driver B with an alternative process possessing the attributes to capture the intended stylized facts. Perhaps the most common option utilized in the literature is *fractional Brownian motion* $B^H = \{B^H(t), t \geq 0\}$, defined as a Gaussian process with $B^H(0) = 0$ a.s., $\mathbb{E}[B^H(t)] = 0$ for all $t \geq 0$ and

$$
\mathbb{E}\left[B^H(t)B^H(s)\right] = \frac{1}{2}\left(t^{2H} + s^{2H} - |t-s|^{2H}\right), \quad s, t \geq 0,
\tag{23}
$$

where H can take values in $(0,1)$ and is called the *Hurst index* (this parameter is named after Harold Edwin Hurst (1880–1978) who studied long-range dependence in fluctuations of the water level in the River Nile). Fractional Brownian motion was initially considered by Kolmogorov [127] and later reintroduced by Mandelbrot and van Ness [128], who also obtained its Volterra-type representation

$$
B^H(t) := \frac{1}{\Gamma\left(H + \frac{1}{2}\right)}\int_{-\infty}^{0}\left((t-s)^{H-\frac{1}{2}} - (-s)^{H-\frac{1}{2}}\right)dB(s) + \frac{1}{\Gamma\left(H + \frac{1}{2}\right)}\int_{0}^{t}(t-s)^{H-\frac{1}{2}}dB(s),
\tag{24}
$$

where $B = \{B(t),\ t \geq 0\}$ is a standard Brownian motion. In the literature, a truncated version of Equation (24) called the *Riemann–Liouville fractional Brownian motion* is often used:

$$B_{RL}^H(t) := \frac{1}{\Gamma\left(H + \frac{1}{2}\right)} \int_0^t (t - s)^{H - \frac{1}{2}} dB(s), \tag{25}$$

and the Volterra kernel

$$\mathcal{K}(t,s) := \frac{1}{\Gamma\left(H + \frac{1}{2}\right)} (t - s)^{H - \frac{1}{2}} \mathbb{1}_{s < t} \tag{26}$$

in Equation (25) is called the *fractional* kernel.

Fractional Brownian motion is extremely convenient from the mathematical perspective: it is the only stochastic process that simultaneously [129]

- is Gaussian;
- has stationary increments;
- is *self-similar*, i.e.,

$$B^H(at) \stackrel{\text{Law}}{=} a^H B^H(t), \quad \forall a \geq 0. \tag{27}$$

In addition, if $H = 1/2$, B^H coincides with a standard Brownian motion and, hence, can be viewed as a broad generalization of the latter. Interestingly, the value $H = 1/2$ turns out to be the boundary between two distinct volatility modeling paradigms, one favoring $H > 1/2$ and the other advocating for $H < 1/2$, and the aim of this subsection is to characterize both of them.

3.3.1. Fractional Models with Long Memory: $H > 1/2$

First of all, observe that

$$\mathbb{E}\left[B^H(1)\left(B^H(n) - B^H(n - 1)\right)\right] = \frac{1}{2}\left(n^{2H} - 2(n - 1)^{2H} + (n - 2)^{2H}\right)$$
$$\sim H(2H - 1)n^{2H - 2}, \quad n \to \infty. \tag{28}$$

Therefore, if $H \in (1/2, 1)$, the autocorrelation function behaves as $O(n^{-\beta})$ with $\beta \in (0, 1)$, revealing the long memory property. In particular, if $H \in (3/4, 1)$, the behavior of Equation (28) matches the empirical estimates $0 < \beta \leq \frac{1}{2}$ for absolute log-returns highlighted in Section 2.3.

Historically, the long memory of fractional Brownian motion for $H > 1/2$ was the original reason to employ the latter in volatility modeling. Namely, in 1998, Comte and Renault [47] suggested the first continuous time fractional volatility model of the form

$$\sigma(t) = \theta_1 \exp\left\{\theta_2 \int_0^t e^{-\theta_3(t - s)} dB^H(s)\right\}, \tag{29}$$

which can mimic volatility persistence and explain slow decays in the smile amplitude of implied volatility surfaces when $T \to \infty$ (in this regard, we also recommend the simulation study [48] that illustrates this phenomenon numerically). Other contributions studying stochastic volatility driven by fractional Brownian motion with $H > 1/2$ include:

- Ref. [130], which considers the model of the form

$$dS(t) = \sigma(t)dW(t),$$
$$\sigma(t) = F\left(\int_0^t a(t, u)dB^H(u) + f(t)\xi_0\right)$$

where F, a, and f are nuisance parameters and ξ_0 is a random initial condition;

- Refs. [131,132], which discuss a fractional counterpart of the model Equation (16):

$$dS(t) = \mu S(t)dt + \sigma(Y(t))S(t)dW(t),$$
$$Y(t) = -\theta_1 Y(t)dt + \theta_2 dB^H(t),$$

(30)

where $\mu, \theta_1, \theta_2 > 0$, and σ is a deterministic function that is additionally assumed to have sublinear growth in [132];
- the series of papers [133–135] that proposes

$$dS(t) = \mu S(t)dt + \sigma(Y(t))S(t)dW(t),$$
$$dY(t) = \frac{1}{2}\left(\frac{\theta_1}{Y(t)} - \theta_2 Y(t)\right)dt + \frac{\theta_3}{2}dB^H(t),$$

(31)

where $\mu, \theta_1, \theta_2, \theta_3 > 0$, and σ is a given function with sublinear growth. Note that Equation (31) can be regarded as a fractional extension of the Heston model Equation (17) since the process $X(t) = Y^2(t)$ satisfies the pathwise SDE

$$dX(t) = (\theta_1 - \theta_2 X(t))dt + \theta_3\sqrt{X(t)}dB^H(t).$$

In some models, long memory is incorporated not directly through fractional Brownian motion, but rather by combining standard Brownian diffusions with the fractional kernel Equation (26). Examples of this approach can be found in, e.g., [131,136], where the prices are assumed to follow

$$dS(t) = \mu(t)S(t)dt + \sigma(t)S(t)dW(t),$$
$$\sigma^2(t) = \theta + \frac{1}{\Gamma\left(H + \frac{1}{2}\right)} \int_{-\infty}^{t} (t - s)^{H - \frac{1}{2}} X(s)ds,$$
$$dX(t) = \theta_1(\theta_2 - X(t))dt + \theta_3\sqrt{X(t)}dB(t),$$

where W and B are correlated standard Brownian motions.

As a final remark, we mention the special interplay between the long memory and self-similarity Equation (27) properties of fractional Brownian motion. As previously discussed in Section 2.3, any measures of long-range dependence are hard to estimate statistically: the analysis of time series during overly extended time periods naturally raises concerns about data non-stationarity. However, if the data are additionally assumed to display self-similarity, its long-term behavior can be inferred from high-frequency observations over shorter periods. For a detailed discussion on self-similarity in financial time series, as well as its connection with long memory, we refer the reader to [1] subsections 2.3 and 2.4.

3.3.2. Rough Revolution: $H < 1/2$

As highlighted in the simulation study [48], fractional Brownian motion with $H > 1/2$ indeed allows the behavior of implied volatility surfaces to be captured for longer maturities. However, a high Hurst index does not seem to have any positive impact on the short-term fit. For example, Alòs, León, and Vives in section 7.2.1 of their 2007 paper [126] analyze the behavior of the short-term implied volatility skew, Equation (9), generated by a variation of the model Equation (30) with $H > 1/2$, and analytically prove that it does not behave as $O(T^{-\beta})$, $\beta \approx \frac{1}{2}$, when $T \to 0$. Interestingly, in section 7.2.2 of the same paper, they notice that the stochastic volatility model driven by a Riemann–Liouville fractional Brownian motion with $H \in (0, 1/2)$ produces the required skew asymptotics $O(T^{-\frac{1}{2}+H})$ as $T \to 0$.

The arguments in [126] were based on *Malliavin calculus* (see, e.g., [137,138]): under certain regularity assumptions, the short-term explosion of the implied volatility skew translates to the explosion of the Malliavin derivative of volatility, the property which holds for, e.g., fractional Brownian motion with $H < 1/2$. In 2014 (although [52] appeared in

Quantitative Finance in 2018, the preprint had been available on SSRN and ArXiv since 2014), Gatheral, Jaisson, and Rosenbaum [52] advocated for $H < 1/2$ from a different perspective: using estimation techniques based on power variations, they came to a conclusion that volatility must have Hölder regularity of order ≈ 0.1. Using Equation (23), it is easy to check that

$$\mathbb{E}[|B^H(t) - B^H(s)|^2] = |t - s|^{2H}, \quad s, t \geq 0,$$

so, by the Kolmogorov–Chentsov theorem (see, e.g., [139] p. 192), the paths of B^H are Hölder continuous up to the order H. In other words, ref. [52] concluded that fractional Brownian motions with very small Hurst index H are preferable choices for stochastic volatility modeling. Later, in 2021, Fukasawa [140] re-visited the interplay between roughness and the power law of the at-the-money volatility skew. The main theoretical result presented in [140] is as follows: *if* the price S is a positive *continuous* semimartingale and *if* the at-the-money implied volatility skew, Equation (9), exhibits the power-law behavior of the form $T^{-\frac{1}{2}+H}$, $T \to 0$, $H \approx 0$, then H_0-Hölder continuity of the quadratic variation derivative

$$\frac{d}{dt}\langle \log S \rangle_t \tag{32}$$

leads to an arbitrage opportunity if $H_0 > H$. Note that Equation (32) exactly coincides with $\sigma = \{\sigma(t), \ t \geq 0\}$ in the general stochastic volatility model Equation (11) and, hence, ref. [140] implies that, in a continuous setting with the power law of the implied volatility skew, the volatility process "*has to*" be rough to avoid arbitrage.

These collective observations rapidly developed into a vast research field known as "*rough volatility*", with hundreds of papers published over the years. For an extensive literature list on the topic that is regularly updated by specialists in the field, we refer the reader to [141]. Some notable models include (in all cases, $H \in \left(0, \frac{1}{2}\right)$):

- the rough fractional stochastic volatility (RFSV) model [52]

$$\sigma(t) = \exp\left\{\theta_1 + \theta_2 \int_{-\infty}^{t} e^{-\theta_3(t-s)} dB^H(s)\right\},$$

which can be regarded as a rough counterpart of the model Equation (29);
- the rough Bergomi model [142,143]

$$dS(t) = \sqrt{\sigma(t)}S(t)dW(t),$$
$$\sigma(t) = \theta_0(t)\exp\left\{2\theta_1 \int_0^t (t-s)^{H-\frac{1}{2}} dB(s) - \theta_1^2 t^{2H}\right\}, \tag{33}$$

where $\theta_1 > 0$ and θ_0 is a deterministic function;
- the mixed rough Bergomi model [144,145]

$$dS(t) = S(t)\sqrt{\sigma(t)}dB(t),$$
$$\sigma(t) = \theta_0(t)\left(\delta \exp\left\{2\theta_1 \int_0^t (t-s)^{H-\frac{1}{2}} dW(s) - \theta_1^2 t^{2H}\right\}\right.$$
$$\left. + (1-\delta) \exp\left\{2\theta_2 \int_0^t (t-s)^{H-\frac{1}{2}} dW(s) - \theta_2^2 t^{2H}\right\}\right), \tag{34}$$

where $\delta \in [0, 1]$, $\theta_1, \theta_2 > 0$, and θ_0 is a deterministic function;

- the rough SABR model [146] where the price is given by

$$dS(t) = \sqrt{\sigma(t)}\beta(S(t))dW(t),$$

$$\sigma(t) = \theta_0(t)\exp\left\{\theta\sqrt{2H}\int_0^t (t-u)^{H-\frac{1}{2}}dB(s) - \frac{1}{2}\theta^2 t^{2H}\right\}dB(t),$$

with $0 \le t \le s$, θ_0 being a positive random process and β being a positive continuous function;

- the rough Stein–Stein model [147]

$$dS(t) = \theta\sigma(t)S(t)dW(t),$$

$$\sigma(t) = \frac{1}{\Gamma\left(H+\frac{1}{2}\right)}\int_0^t (t-s)^{H-\frac{1}{2}}dB(s);$$

- the fast-varying rough volatility [148]

$$\begin{aligned}
dS(t) &= \mu(t)dt + \sigma^\varepsilon(t)dW(t),\\
\sigma^\varepsilon(t) &= F(Z^\varepsilon(t)),\\
Z^\varepsilon(t) &= \frac{\theta}{\sqrt{\varepsilon}}\int_{-\infty}^t \mathcal{K}_H\left(\frac{t-s}{\varepsilon}\right)dB(s),
\end{aligned} \tag{35}$$

where $\mathcal{K}_H(t) := C_H\left(t^{H-\frac{1}{2}} - \int_0^t (t-s)^{H-\frac{1}{2}}e^{-s}ds\right)$, $C_H > 0$, F is positive and bounded one-to-one function, $\theta > 0$, and $\varepsilon > 0$ is assumed to be small;

- the rough Heston model [149]

$$\begin{aligned}
dS(t) &= \sqrt{\sigma(t)}S(t)dW(t),\\
\sigma(t) &= \sigma(0) + \int_0^t \frac{(t-s)^{H-\frac{1}{2}}}{\Gamma\left(H+\frac{1}{2}\right)}\left(\theta_1(\theta_2 - \sigma(s))ds + \theta_3\sqrt{\sigma(s)}dB(s)\right)
\end{aligned} \tag{36}$$

and its modification, the quadratic rough Heston model [150,151],

$$\begin{aligned}
dS(t) &= S(t)\sqrt{\sigma(t)}dW(t),\\
\sigma(t) &= a(Z(t)-b)^2 + c,\\
Z(t) = \int_0^t \theta_1\frac{(t-s)^{H-\frac{1}{2}}}{\Gamma\left(H+\frac{1}{2}\right)}(\theta_0(s) - Z(s))ds &+ \int_0^t \theta_2\frac{(t-s)^{H-\frac{1}{2}}}{\Gamma\left(H+\frac{1}{2}\right)}\sqrt{a(Z(s)-b)^2 + c}dW(s),
\end{aligned} \tag{37}$$

where B, W are Brownian motions, $a, b, c, \theta_1, \theta_2, \theta_3 > 0$, and θ_0 is a deterministic function.

The rough Heston approach Equation (36) is especially interesting from the modeling perspective: in addition to the power-law behavior of Equation (9), it can reproduce a special form of the Zumbach effect [152,153] and can be interpreted as the limit of a reasonable tick-by-tick price model based on two-dimensional Hawkes processes [154], i.e., it deduces the roughness of the volatility directly from the market microstructure.

3.3.3. Puzzles of Fractionality and Roughness

In principle, one can divide the majority of arguments in favor of rough volatility into two distinct categories:

- econometric studies like [52], which involve regularity estimations of volatility from historical high-frequency samples of an asset (e.g., S&P500 index);

- analyses from the options pricing perspective such as [126,140] that justify roughness by its ability to reproduce, e.g., the power-law behavior of implied volatility skews, Equation (9).

However, as is often the case when modeling extremely complex systems like the financial market, neither of these arguments is convincing enough to be unequivocally declared undeniable and immune to valid criticism. For example, the methodology used in [52] was examined by Rogers [155], Cont and Das [156], as well as Fukasawa, Takabatake, and Westphal [157,158]: they apply the same regularity estimation procedure as in [52] to *synthetic* datasets and report that the estimator tends to produce low values of H *regardless of the true parameter* used in the simulation. This issue is explained as follows in the above-mentioned literature: since the volatility σ is not observable directly, one must first extract some volatility proxy from the stock data and then use *it* as the sample for regularity estimation. However, this procedure results in additional approximation errors that bias the estimator of H towards zero. The authors of [52] acknowledge this issue themselves, writing that "*estimation errors when estimating volatility can be quite significant for some models, leading to downward biases in the measurement of the smoothness*" [52] p. 946.

Interestingly, there were several subsequent attempts to account for the problem described above. For example, Fukasawa, Takabatake, and Westphal [157,158] develop a more robust estimation technique and still report that $H < 0.1$ is indeed the best fit for the volatility model

$$d \log \sigma^2(t) = \theta_1(t)dt + \theta_2 dB^H(t) \tag{38}$$

with θ_1 being an unknown adapted càdlàg process. Bolko et al. [159] perform the generalized method of moment (GMM) estimation and also confirm the roughness of volatility for thirty-one leading stock indexes under the assumption that $\log \sigma^2$ is a fractional Ornstein–Uhlenbeck process.

Next, the claim that the implied volatility skew follows the power law can also be challenged. For example, Guyon and El Amrani [49] investigate the volatility skew behavior in more detail and conclude that

> "*...power law fits the term-structure of ATM skew of equity indexes quite well over a large range of maturities, from 1 month to a few years. But this should not lead one to conclude that the ATM skew blows up like this power law at zero maturity. Zooming in on the short maturities shows that the power-law extrapolation fits the data poorly.*"

In summary, Guyon and El Amrani ([49] p. 14) argue that "*far from being infinite, the zero-maturity extrapolation of the skew is distributed around a typical value of 1.5*" (in this context, "*the skew*" means the value $\Psi(0)$ in Equation (9)). On the contrary, Delemotte, De Marco, and Ségonne [50] investigate the behavior of average volatility skews between 2007 and 2015 (a period that is different from the period 2020–2022 of Guyon and El Amrani [49]) and find that a power-law-type explosive behavior of the implied volatility skew when $T \to 0$ is entirely appropriate! In addition, they argue ([50] section 3) that the average skew behaves differently for shorter and longer maturities. Their analysis suggests that

$$\mathbb{E}[\Psi(T)] \propto \begin{cases} T^{-\frac{1}{2}+H_1}, & T < T_1, \\ T^{-\frac{1}{2}+H_2}, & T > T_1, \end{cases} \tag{39}$$

where $H_1 > H_2$ and T_1 is roughly 2 months.

What is the reason for such a discrepancy in conclusions regarding very similar (and sometimes *literally the same*) datasets? Features like the behavior of the volatility skew are very intricate and their assessment based on discrete data leaves much space for interpretation and extrapolation. For example, Delemotte, De Marco, and Ségonne ([50] pp. 2–3) compare their result to [49] and summarize the difference as follows:

> "*As a result, both models could very reasonably be used to extrapolate the ATM skew for very short maturities, and in any case below the first maturity quoted on the market, while leading to different skew asymptotics: finite limiting skew in the 2-factors Bergomi model,*"

> *and exploding skew in the 2PL model—indicating that* **the question of the explosive or non-explosive nature of the short-end of the skew curve might be (at least in the case of the SP500 index) hard to disambiguate***.*"

In some sense, this ambiguity is very natural: the volatility itself is a purely theoretical (albeit meaningful) concept, and hence there is no single "*ultimately true volatility model*". In this context, the question "*is volatility **truly** rough*?" lacks a definitive "*correct*" answer. Rough volatility is a valid modeling framework with its advantages and limitations which indeed mimics some important features of the data; for example, the empirically observed steepness of smiles at-the-money does increase close to maturity (as clearly visible in Figure 2), and rough models seem to capture this effect better than classical Brownian diffusions, irrespective of whether this increase "truly" follows the power law. In turn, econometric analyses like [157–159] can be viewed as reasonable reality checks which can reveal possible points of contradiction with data and test specific models against each other; for example, the results of [157–159] could mean that Equation (38) with small H seems to be more consistent with some features of the data than, e.g., Equation (18) with $\beta = 1$ or Equation (29) with $H > \frac{1}{2}$. On the other hand, other models may also be compatible with observations. After all, according to a famous aphorism commonly attributed to the statistician George Box, *all models are wrong, but some are useful*.

With this disclaimer in mind, let us finish this subsection with a discussion of another intricate point coming from the interplay between roughness and long memory *in the specific context of fractional Brownian motion B^H*. More precisely, *long memory* requires the Hurst index $H > 1/2$ whereas *roughness* demands $H < 1/2$. Despite some studies that demonstrate spurious long memory appearing due to model mis-specifications (see, e.g., [52] section 4), roughness alone cannot explain the behavior of the entire volatility surface. For example, Funahashi and Kijima [160] demonstrate that volatility models based on fractional Brownian motion with $H < 1/2$ do not give the required rate of decrease in the smile amplitude as $T \to \infty$, whereas $H > 1/2$ gives this effect (see also [48]).

It should be noted that this contradiction, referred to as the "*fractional modeling puzzle*" in [160], is somewhat synthetic: in general, long memory and roughness do not depend on each other and can co-exist within a single stochastic process. In other words, the reason for this "*fractional puzzle*" comes from the properties of fractional Brownian motion itself and a different modeling framework may utilize both of these features without any structural contradictions (see, e.g., the model based on Brownian semi-stationary processes by Bennedsen, Lunde, and Pakkanen [161]). Nevertheless, in a continuous setting, B^H seems to be an extremely convenient and valuable asset from the mathematical perspective:

- it is Gaussian, enabling the utilization of numerous methods from Gaussian process theory including efficient numerical methods, Malliavin calculus, etc. (see, e.g., [138] chapter 5);
- it has stationary increments and is ergodic, which facilitates various statistical estimation techniques (see, e.g., [162]);
- despite being non-semimartingale, fractional Brownian motion has a developed stochastic integration theory [129].

In the literature, there are several possible approaches to incorporate long memory and roughness within the fractional framework. The most straightforward methodology is to use two fractional Brownian motions with different Hurst indices. Such a model was utilized in, e.g., [160] (see also [137] section 7.7) in the form

$$
\begin{aligned}
dS(t) &= \mu S(t)dt + \sigma(X^1(t), X^2(t))dW(t), \\
dX^i(t) &= (\theta_1^i - \theta_2^i X^i(t))dt + \theta_3^i dB^{H_i}(t), \quad i = 1, 2,
\end{aligned}
\tag{40}
$$

with B^{H_1}, B^{H_2} being two fractional Brownian motions with $H_1 > 1/2$ and $H_2 < 1/2$. The authors report that this model indeed manages to grasp the implied volatility surface with the power law for short maturities and slower decay of the smile amplitude for long maturities.

Such a suggestion somehow resonates with the observations made in [50]: the authors also discuss the possibility of introducing several factors with different regularity into the model in order to reproduce Equation (39) (although they consider $0 < H_2 < H_2 < \frac{1}{2}$).

Another interesting possibility—*a multifractional Brownian motion*—is advocated in Corlay et al. [163] (see also [164]). There, the authors estimate the local roughness of the volatility and conclude that it is heavily variable and has periods of low (≈ 0.1) and high (≈ 0.8) regularity (see [163] figure 2). It is also important to note that [52] section 2.6 also reports some dependence of the volatility roughness on time.

As a generalization of both approaches mentioned above, one can also consider the usage of *Gaussian Volterra processes* $Z(t) := \int_0^t \mathcal{K}(t,s)dB(s)$. Such volatility drivers were considered in, e.g., [165,166] or the series of papers [167,168].

3.4. Usability Challenges of Stochastic Volatility Models

We finish this section with several remarks that, in our opinion, are worth the attention of the reader.

3.4.1. Positivity of Volatility

As mentioned in Section 3.2, one of the natural expectations from a "*reasonable*" volatility model is the positivity of its paths. This requirement actually goes beyond the simple consideration that $\sigma = \{\sigma(t), t \geq 0\}$ should resemble its original proxy, Equation (1), and is connected to the procedure of transition between the physical and the pricing measures. In the Black–Scholes-type models Equation (11), martingale densities usually involve terms of the form $\int_0^T \frac{1}{\sigma(s)}dW(s)$ and $\int_0^T \frac{1}{\sigma^2(s)}ds$ (see, e.g., [169] proposition 1.11) which can be poorly defined if the volatility hits zero with positive probability. Of course, one may model the market under the pricing measure in the first place, but, in this situation, one sacrifices the ambition to justify the approach with econometric analysis based on historical time series as in [52] which, of course, should be performed under the physical measure.

3.4.2. Possibility of Moment Explosions

A common issue for the stochastic volatility framework is the possibility of moment explosions in price [99]. That means that the moment of the price $\mathbb{E}[S^r(t)]$ may be infinite for all time points t after some t_*. Moment explosions can be a notable drawback from at least two perspectives. First, numerical schemes involving L^r-convergence become inaccessible. Another perspective is asset pricing, as noted in [99] section 8, "*several actively traded fixed-income derivatives require at least L^2 solutions to avoid infinite model prices*". In principle, both positivity and the absence of moment explosions can be achieved by assuming bounds on the volatility; see, e.g., [123,132,148,167,168,170,171] and the footnote on p. 2 in [151]. It seems that this assumption is not too aggravating: in industry, there seems to be a consensus [172] that real-world proxies of volatility are typically range-bound.

3.4.3. Importance of Numerical Methods

As noted in [18] p. 24, the standard Heston model is still widely used despite all the empirical inconsistencies outlined in Section 3.2. The reason for this is mainly the availability of algorithms for practically all possible applications. Stochastic volatility models normally do not have closed-form expressions for option pricing, portfolio optimization, hedging, etc., and, therefore, the development of efficient numerics for them is of acute importance. It is especially relevant for fractional/rough models which are predominantly non-Markovian and, hence, cannot utilize various stochastic optimization techniques developed for Markov processes. In recent years, some work in this direction has been performed in, e.g., [151,167,168,173–175].

4. VIX and the Joint Calibration Puzzle

As previously mentioned in the Introduction, volatility is an important tool for quantifying risk, in addition to being a critical variable in option pricing. Therefore, it is no surprise that market actors have been seeking dedicated financial instruments to hedge

against drastic volatility changes and capitalize on overall volatility trends. Nowadays, derivatives written on various volatility proxies are extremely popular among investors: for instance, according to the Cboe Annual Report 2022 [176], the *CBOE Volatility Index* (*VIX*, see Figure 6) is one of the most traded underlying assets of the Chicago Board Options Exchange, alongside the S&P 500 (SPX) index.

Since volatility is not a directly observable value and, in some sense, it comes from the theoretical domain, its quantification is a necessary step before considering its use as an underlying asset. Naturally, various methodologies have been applied for such quantification; a detailed historical overview of this subject can be found in, e.g., [177] section 3. Here, we mention the early approaches of Gastineau (1977) [178] and Cox and Rubinstein (1985) [179] appendix 8A, who proposed indices based on the implied volatility, as well as Brenner and Galai (1989) [180], who suggested a metric coming from the realized volatility and considered derivatives employing it as the underlying asset. Another index utilizing an average of S&P100 option implied volatilities was introduced by Fleming et al. [181], and this methodology was used by the Chicago Board Options Exchange for computation of their early version of VIX between 1993 and 2003. The general idea behind the modern computation of VIX can be traced back to Breeden and Litzenberger (1978) [182], but the exact formulas were crystallized in the works of Dupire [183] and Carr and Madan [184].

Nowadays, VIX is the preeminent volatility proxy on a global scale, closely connected to a volatility modeling challenge occasionally referred to as "*The Holy Grail of volatility modeling*" [150,185]. In order to outline this problem in more detail and prevent any ambiguity, let us first dedicate some time and delve into the specifics of VIX computation and interpretation.
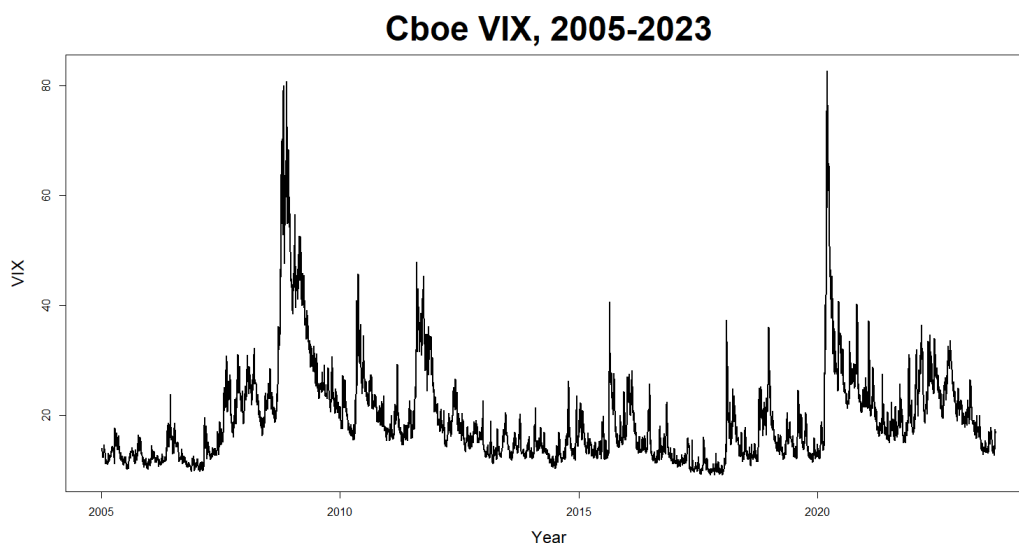
Figure 6. Daily values of the CBOE VIX index, 2005–2023. Note the two spikes in 2008 and 2020 which correspond to the respective economic crises. The data were retrieved from CBOE.com [186].

4.1. Intermezzo: VIX and Its Interpretation

Assume that the SPX forward price *S* is a continuous martingale satisfying

$$dS(t) = \sigma(t)S(t)dW(t), \tag{41}$$

on a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{Q})$ under the risk-neutral probability measure $\mathbb{Q}$, where *W* is a $\mathbb{Q}$-Brownian motion and $\sigma = \{\sigma(t), \, t \geq 0\}$ is an adapted square integrable

stochastic process. It is easy to verify that, under some mild assumptions on σ, Equation (41) has a unique solution of the form

$$S(t) = S(0) \exp\left\{ -\frac{1}{2} \int_0^t \sigma^2(s)ds + \int_0^t \sigma(s)dW(s) \right\}.$$

Next, fix a period of time T (in the case of CBOE VIX, $T = 30$ days) and consider the value

$$-\frac{2}{T}\mathbb{E}_{\mathbb{Q}}\left[\log \frac{S(t+T)}{S(t)} \,\Big|\, \mathcal{F}_t \right]. \tag{42}$$

One the one hand, since, by the martingale property,

$$\mathbb{E}_{\mathbb{Q}}\left[\int_t^{t+T} \sigma(s)dW(s) \,\Big|\, \mathcal{F}_t \right] = 0,$$

one can observe that

$$
\begin{aligned}
-\frac{2}{T}\mathbb{E}_{\mathbb{Q}}\left[\log \frac{S(t+T)}{S(t)} \,\Big|\, \mathcal{F}_t \right] &= -\frac{2}{T}\mathbb{E}_{\mathbb{Q}}\left[-\frac{1}{2}\int_t^{t+T}\sigma^2(s)ds + \int_t^{t+T}\sigma(s)dW(s) \,\Big|\, \mathcal{F}_t \right] \\
&= \frac{1}{T}\int_t^{t+T}\mathbb{E}_{\mathbb{Q}}\left[\sigma^2(s) \,\big|\, \mathcal{F}_t \right]ds.
\end{aligned}
\tag{43}
$$

On the other hand, by Taylor's formula with integral remainder,

$$
\begin{aligned}
\log \frac{S(t+T)}{S(t)} &= \frac{S(t+T)-S(t)}{S(t)} - \int_{S(t)}^{S(t+T)}\frac{(S(t+T)-K)}{K^2}dK \\
&= \frac{S(t+T)-S(t)}{S(t)} - \int_0^{S(t)}\frac{(K-S(t+T))_+}{K^2}dK - \int_{S(t)}^{\infty}\frac{(S(t+T)-K)_+}{K^2}dK,
\end{aligned}
$$

and hence, since

$$\mathbb{E}_{\mathbb{Q}}\left[\frac{S(t+T)-S(t)}{S(t)} \,\Big|\, \mathcal{F}_t \right] = 0$$

by the martingale property, Equation (42) can be re-written as

$$
\begin{aligned}
-\frac{2}{T}\mathbb{E}_{\mathbb{Q}}\left[\log \frac{S(t+T)}{S(t)} \,\Big|\, \mathcal{F}_t \right] &= \mathbb{E}_{\mathbb{Q}}\left[\frac{2}{T}\int_0^{S(t)}\frac{(K-S(t+T))_+}{K^2}dK \,\Big|\, \mathcal{F}_t \right] \\
&\quad + \mathbb{E}_{\mathbb{Q}}\left[\int_{S(t)}^{\infty}\frac{(S(t+T)-K)_+}{K^2}dK \,\Big|\, \mathcal{F}_t \right] \\
&= \frac{2}{T}\int_0^{S(t)}\frac{\mathbb{E}_{\mathbb{Q}}[(K-S(t+T))_+ \,|\, \mathcal{F}_t]}{K^2}dK \\
&\quad + \int_{S(t)}^{\infty}\frac{\mathbb{E}_{\mathbb{Q}}[(S(t+T)-K)_+ \,|\, \mathcal{F}_t]}{K^2}dK \\
&= \frac{2e^{rT}}{T}\left(\int_0^{S(t)}\frac{P_t(K,T)}{K^2}dK + \int_{S(t)}^{\infty}\frac{C_t(K,T)}{K^2}dK \right),
\end{aligned}
\tag{44}
$$

where r is the instantaneous interest rate and $P_t(K,T)$, $C_t(K,T)$ are, respectively, the prices (at moment t) of put and call options with strike K and expiry date $t+T$. To summarize, under the settings specified above, Equations (43) and (44) yield

$$\frac{1}{T}\int_t^{t+T}\mathbb{E}_{\mathbb{Q}}\left[\sigma^2(s) \,\big|\, \mathcal{F}_t \right]ds = \frac{2e^{rT}}{T}\left(\int_0^{S(t)}\frac{P_t(K,T)}{K^2}dK + \int_{S(t)}^{\infty}\frac{C_t(K,T)}{K^2}dK \right). \tag{45}$$

VIX is computed (see, e.g., [187] or [188] section 1.5.2) by discretizing the right-hand side of Equation (45) (note that the latter relies exclusively on the empirically observable values, SPX forward $S(t)$ and put/call option prices):

$$\text{VIX}_t^2(T) := \frac{2e^{rT}}{T}\left(\sum_{j=-M}^{-1}\frac{P_t(K_j,T)}{K_j^2}(K_{j+1}-K_j) + \sum_{i=1}^{N}\frac{C_t(K_i,T)}{K_i^2}(K_i-K_{i-1})\right) - \frac{1}{T}\left(\frac{S(t)}{K_0}-1\right)^2,$$

where $K_{-M} < K_{-(M-1)} < ... < K_0$ are payoffs of all listed out-of-the-money put options with maturity at $t+T$, $K_0 \leq S(t)$, and $K_1 < K_2 < ... < K_N$ are payoffs of all listed out-of-the-money call options with maturity at $t+T$, $K_1 > S(t)$. In the meanwhile, the left-hand side of Equation (45) is a standard notion of the forward variance, precisely the value one intends to capture.

4.2. VIX Smile and the Joint Calibration Puzzle

Despite any potential inaccuracies, VIX stands as the most popular proxy for market volatility in the world and, therefore, it is unsurprising that derivatives with VIX underlying them are actively traded on the market. In particular, in March 2004, CBOE launched VIX futures and, in February 2006, VIX options were introduced—and, naturally, a problem of pricing such derivatives emerged.

First of all, having VIX option prices, one can use the procedure described in Section 2.2 and construct the corresponding implied volatility surface $\widehat{\sigma}(T,K)$, with T denoting the time to maturity and K being the strike. Contrary to the convex "smiles" for derivatives written on classical stocks and stock indices, the VIX implied volatility $\widehat{\sigma}(T,K)$ consistently exhibits concave behavior in K for every fixed T with positive slope at the money (see [189] figure 1, as well as [119,124], or [190]).

It turns out that continuous stochastic volatility models have mixed success in reproducing this phenomenon, especially for shorter maturities—in this regard, we recommend an excellent review [191] on the topic. Alòs, García-Lorite, and Gonzalez [189] (see also [137] chapter 10) use Malliavin calculus tools and analyze selected models to check whether they produce the positive slope at the money for short maturities. Their results show that:

- The SABR model, Equation (19), gives a flat skew, but its modification, called the *mixed SABR model*

$$\sigma(t) = \theta_0\left(\delta\exp\left\{2\theta_1 W(t) - \theta_1^2 t\right\} + (1-\delta)\exp\left\{2\theta_2 W(t) - \theta_2^2 t\right\}\right),$$

  where $\delta \in [0,1]$, θ_0, θ_1, $\theta_2 > 0$, generates the required positive slope provided that $\delta \neq 0, 1$ and $\theta_1 \neq \theta_2$;
- The Heston model, Equation (17), normally generates negative VIX skew (see also [113] Figure 3); however, under some conditions on coefficients (e.g., when Feller condition is violated, see [190]), the slope can be positive;
- The rough Bergomi model, Equation (33), as well as its mixed modification, Equation (34), both give the required slope (see also [143]).

Other notable continuous stochastic volatilities that are able to capture the upward smirk include:

- The 3/2 model, Equation (20) [119] section 2;
- The double CEV model, Equation (21) [18];
- The Heston with stochastic vol-of-vol model, Equation (22) [124];
- The quintic Ornstein–Uhlenbeck model [192]:

$$\sigma(t) = \sqrt{\theta_0(t)}\frac{p(X(t))}{\sqrt{\mathbb{E}[p^2(X(t))]}},$$

$$X(t) = \varepsilon^{-\theta}\int_0^t e^{-\frac{\theta}{\varepsilon}(t-s)}dB(s),$$

(46)

and, more generally, a class of Gaussian polynomial models [174]:

$$\sigma(t) = \sqrt{\theta_0(t)} \frac{p(X(t))}{\sqrt{\mathbb{E}[p^2(X(t))]}},$$

$$X(t) = \int_0^t \mathcal{K}(t-s)dB(s), \tag{47}$$

where $\theta > 0$, p is a polynomial, and $\mathcal{K}$ is a square integrable Volterra kernel;

- The quadratic rough Heston model, Equation (37).

As one can see, multiple continuous models, both standard Brownian diffusions and rough volatility models, are able to reproduce the shape of the VIX smile. However, there is an additional dimension to that problem: since there is a direct connection between SPX and VIX, models calibrated with respect to the SPX data should be consistent with VIX and vice versa. In other words, one seeks a market model that would *jointly calibrate to both VIX and SPX simultaneously*. As noted by Guyon in [185], "*Without such models, financial institutions could possibly arbitrage each other, and even market making desks within the same institution could do so...*".

It turns out that this problem, known as the "*joint calibration puzzle*", is extremely difficult to solve; so much so that some researchers label it "*the Holy Grail of volatility modeling*" [150,185]. For a long time, continuous-time stochastic volatility models without jumps failed to produce the perfect SPX-VIX joint calibration. Some notable attempts include the 3/2 model [119], double CEV model [18], Heston with stochastic vol-of-vol model [124], and rough Bergomi model [143]—in all cases, the joint fit was only partially successful, predominantly failing for short maturities (see also an excellent presentation by J. Guyon [144] on the topic). The most fruitful approaches in solving the joint calibration puzzle were outside of the continuous-time continuous-price paradigm:

- Cont and Kokholm [193] use jump diffusions to directly model forward variances and SPX; jumps indeed allow the ATM S&P500 skew and the ATM VIX implied volatility to be de-coupled, giving a good joint fit;
- Guyon [185,194] utilizes nonparametric discrete time model building joint probability measures on SPX and VIX on a discrete set of dates; this allows an almost perfect joint calibration to be obtained.

However, in recent years there have been several successful attempts inside the fully continuous framework, all with a higher number of parameters:

- Gatheral, Jusselin, and Rosenbaum [150] and Rosenbaum and Zhang [151] show that the quadratic rough Heston model, Equation (37), calibrates very well to SPX and VIX smiles;
- Abi Jaber, Illand, and Li [174,192] provide a good fit with Gaussian polynomial models, Equation (47);
- Bourgey and Guyon ([195], Section 4) extend the discrete time model of [185,194] to a continuous-time setting;
- Guyon and Mustapha [196] achieve an impressive calibration with the model

$$dS_t = -\frac{1}{2}\sigma_S^2(t, S_t, Y_t)dt + \sigma_S(t, S_t, Y_t)dW_t,$$

$$dY_t = \mu_Y(t, S_t, Y_t)dt + \sigma_Y(t, S_t, Y_t)\left(\rho(t, S_t, Y_t)dW_t + \sqrt{1-\rho^2(t, S_t, Y_t)}dB_t\right),$$

where the drift, volatility, and correlation functions σ_S, μ_Y, σ_Y, and ρ are modeled with neural networks.

4.3. VIX and Continuous Models: Interpretation Issues

We finish this section by highlighting a significant caveat related to the interpretation and tractability of VIX within the context of continuous stochastic volatility models.

Despite the undoubted elegance of the VIX computation methodology, the precision of VIX per se in capturing the volatility of SPX heavily depends on the underlying assumptions. Namely, it is crucial that SPX is assumed to follow continuous dynamics, Equation (41) and any deviations from this model lead to some deterioration in the VIX tractability. In particular, if there are jumps in the SPX forward dynamics, Equation (43) does not hold, with an additional term appearing on the right-hand side of the expression (see, e.g., [197]). Therefore, if Equation (41) is violated, Equation (45) does not hold either, and VIX loses its original interpretation.

It is important to note that the viability of the model Equation (41) can be tested statistically by comparing VIX to *variance swap contracts*. By definition, a *variance swap* is a futures contract with a payoff at the moment T of the form

$$\frac{1}{T} \sum_{k=1}^{n} \left(\log \frac{S(t_k)}{S(t_{k-1})} \right)^2 - VS,$$

where $0 = t_0 < t_1 < \cdots < t_n = T$ is a partition of $[0, T]$ and VS denotes the corresponding swap rate that is determined by the market. Non-arbitrage arguments imply that, if the dynamics Equation (41) indeed holds and some mild assumptions on σ are satisfied,

$$VS = \mathbb{E}_{\mathbb{Q}} \left[\frac{1}{T} \sum_{k=1}^{n} \left(\log \frac{S(t_k)}{S(t_{k-1})} \right)^2 \right] \to \frac{1}{T} \int_0^T E_{\mathbb{Q}} \left[\sigma^2(s) \right] ds$$

as $\max_k |t_k - t_{k-1}| \to 0$. In other words, under the model Equation (41), market swap rates $\sqrt{VS}$ should approximately coincide with VIX. However, as reported by Aït-Sahalia et al. in their analysis [197] using the over-the-counter swap rate data, there is a statistically significant gap between VS and VIX^2: the values of $VS - \text{VIX}^2$ are mostly positive, larger during market turmoils but sizable even in quiet times. For more details in that regard, as well as for possible alternatives to the current VIX computation method, we also refer the reader to the discussions in [198–200].

In other words, there is some evidence showing that the tractability of VIX as a volatility index is not that straightforward and the real-world observations of VIX are not sampled from

$$\sqrt{\frac{1}{T} \int_t^{t+T} \mathbb{E}_{\mathbb{Q}} \left[\sigma^2(s) \mid \mathcal{F}_t \right] ds} \tag{48}$$

as models with continuous price dynamics assume (as suggested in [200], in the presence of jumps, VIX measures *the risk-neutral entropy* of the simple return on the S&P 500 index rather than (48)). Therefore, fitting stochastic volatility models with the continuous price dynamics to VIX without any adjustments may in principle produce unstable results.

5. Concluding Remarks

In this survey, we outlined—though with considerably broad brushstrokes—the existing literature on stochastic volatility. Clearly, the number of approaches is very large and we acknowledge that we are far from describing all of them, even within the considerably narrower framework of continuous models. Nevertheless, we hope that the reader received a general understanding of the motivation behind this modeling paradigm.

We conclude our presentation by saying that updating, refining, and improving models still seems an endless race. Right now, new techniques and modeling paradigms are being developed—in particular, we mention [201], which finds strong arguments for fully path-dependent volatility models as well as machine learning and signature methods such as in [202–205] (the model in the latter actually includes the Gaussian polynomial models, Equation (47), as a special case). Perhaps these novel tools will become the new classics and open new frontiers in understanding the financial market.

References

1. Cont, R. Long range dependence in financial markets. In *Fractals in Engineering*; Springer: London, UK, 2005; pp. 159–179.
2. Cutler, D.; Poterba, J.; Summers, L. What moves stock prices? *J. Portf. Manag.* **1989**, *15*, 4–12. [CrossRef]
3. Clark, P.K. A subordinated stochastic process model with finite variance for speculative prices. *Econom. J. Econom. Soc.* **1973**, *41*, 135. [CrossRef]
4. Skiadopoulos, G. Volatility smile consistent option models: A survey. *Int. J. Theor. Appl. Financ.* **2001**, *04*, 403–437. [CrossRef]
5. Shephard, N.; Andersen, T.G. Stochastic volatility: Origins and overview. In *Handbook of Financial Time Series*; Springer: Berlin/Heidelberg, Germany, 2009; pp. 233–254.
6. Duong, D.; Swanson, N.R. Volatility in discrete and continuous-time models: A survey with new evidence on large and small jumps. In *Missing Data Methods: Time-Series Methods and Applications*; Emerald Group Publishing Limited: Bingley, UK, 2011; pp. 179–233.
7. Markowitz, H. Portfolio Selection. *J. Financ.* **1952**, *7*, 77. [CrossRef]
8. Black, F.; Scholes, M. The Pricing of Options and Corporate Liabilities. *J. Political Econ.* **1973**, *81*, 637–654. [CrossRef]
9. Merton, R.C. Theory of rational option pricing. *Bell J. Econ. Manag. Sci.* **1973**, *4*, 141. [CrossRef]
10. Engle, R.F. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econom. J. Econom. Soc.* **1982**, *50*, 987. [CrossRef]
11. Bollerslev, T. Generalized autoregressive conditional heteroskedasticity. *J. Econom.* **1986**, *31*, 307–327. [CrossRef]
12. Nelson, D.B. Conditional heteroskedasticity in asset returns: A new approach. *Econom. J. Econom. Soc.* **1991**, *59*, 347–370. [CrossRef]
13. Duffie, D.; Pan, J.; Singleton, K. Transform analysis and asset pricing for affine jump-diffusions. *Econom. J. Econom. Soc.* **2000**, *68*, 1343–1376. [CrossRef]
14. Barndorff-Nielsen, O.E.; Shephard, N. Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial economics. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **2001**, *63*, 167–241. [CrossRef]
15. Bollerslev, T.; Chou, R.Y.; Kroner, K.F. ARCH modeling in finance. *J. Econom.* **1992**, *52*, 5–59. [CrossRef]
16. Francq, C. *GARCH Models: Structure, Statistical Inference and Financial Applications*, 2nd ed.; Wiley-Blackwell: Hoboken, NJ, USA, 2019.
17. Barndorff-Nielsen, O.E.; Shephard, N. Lévy Driven Volatility Models. 2011. Available online: https://pure.au.dk/ws/files/1938 74249/levybook.pdf (accessed on 30 August 2023).
18. Gatheral, J. *The Volatility Surface: A Practitioner's Guide*; John Wiley & Sons: Nashville, TN, USA, 2006.
19. Rachev, S.T.; Kim, Y.S.; Bianchi, M.L.; Fabozzi, F.J. *Financial Models with Lévy Processes and Volatility Clustering*; John Wiley & Sons: Chichester, UK, 2011.
20. Cont, R.; Tankov, P. *Financial Modelling with Jump Processes*; Chapman & Hall/CRC: Philadelphia, PA, USA, 2003.
21. Shiryaev, A.N. *Essentials of Stochastic Finance: Facts, Models, Theory*; World Scientific Publishing: Singapore, 1999.
22. Mariani, M.C.; Florescu, I. *Quantitative Finance*; John Wiley & Sons: Nashville, TN, USA, 2020.
23. Jovanovic, F.; Le Gall, P. Does God practice a random walk? The "financial physics" of a nineteenth-century forerunner, Jules Regnault. *Eur. J. Hist. Econ. Thought* **2001**, *8*, 332–362. [CrossRef]
24. Regnault, J. *Calcul des Chances et Philosophie de la Bourse*; Mallet Bachelier and Castel: Paris, France, 1863.
25. Bachelier, L. Théorie de la spéculation. *Ann. Sci. l'Ecole Norm. Super. Quatr. Ser.* **1900**, *17*, 21–86. [CrossRef]
26. Bachelier, L.; Davis, M.; Etheridge, A. *Louis Bachelier's Theory of Speculation: The Origins of Modern Finance*; Translated and with Commentary by Mark Davis and Alison Etheridge; Princeton University Press: Princeton, NJ, USA, 2006.
27. Einstein, A. Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen. *Ann. Phys.* **1905**, *322*, 549–560. [CrossRef]
28. Osborne, M.F.M. Brownian Motion in the Stock Market. *Oper. Res.* **1959**, *7*, 145–173. [CrossRef]
29. Samuelson, P.A. Mathematics of speculative price. *SIAM Rev.* **1973**, *15*, 1–42. [CrossRef]

30. Revuz, D.; Yor, M. *Continuous Martingales and Brownian Motion*, 3rd ed.; Springer: Berlin/Heidelberg, Germany, 1999.
31. Boness, A.J. Elements of a Theory of Stock-Option Value. *J. Political Econ.* **1964**, *72*, 163–175. [CrossRef]
32. Sprenkle, C.M. Warrant prices as indicators of expectations and preferences. *Yale Econ. Essays* **1959**, *1*, 179–231.
33. Thorp, E.O.; Kassouf, S.T. *Beat the Market: A Scientific Stock Market System*; Random House: New York, NY, USA, 1967.
34. Samuelson, P.A. Rational theory of warrant pricing. *Ind. Manag. Rev.* **1959**, *6*, 13–39.
35. Samuelson, P.A.; Merton, R.C. A Complete Model of Warrant Pricing that Maximizes Utility. *Ind. Manag. Rev.* **1972**, *10*, 17–46.
36. Black, F. Citation Classic—The Pricing of Options and Corporate Liabilities. *Curr. Contents Soc. Behav. Sci.* **1987**, *33*, 16.
37. Ross, S.A. Finance. In *The New Palgrave Dictionary of Economics*; Eatwell, J., Milgate, M., Newman, P., Eds.; Springer: London, UK, 1987; Volume 2, pp. 322–336.
38. Fengler, M. *Semiparametric Modeling of Implied Volatility*, 2005th ed.; Springer: Berlin/Heidelberg, Germany, 2005.
39. Wiggins, J.B. Option values under stochastic volatility: Theory and empirical estimates. *J. Financ. Econ.* **1987**, *19*, 351–372. [CrossRef]
40. Jackwerth, J.C.; Rubinstein, M. Recovering probability distributions from option prices. *J. Financ.* **1996**, *51*, 1611. [CrossRef]
41. Latané, H.A.; Rendleman, R.J. Standard deviations of stock price ratios implied in option prices. *J. Financ.* **1976**, *31*, 369–381. [CrossRef]
42. Lee, R.W. Implied volatility: Statics, dynamics, and probabilistic interpretation. In *Recent Advances in Applied Probability*; Springer: Boston, MA, USA, 2006; pp. 241–268.
43. S&P 500 Index. Available online: https://finance.yahoo.com/quote/%5ESPX?p=%5ESPX (accessed on 3 May 2023).
44. Derman, E.; Miller, M.B.; Park, D. *The Volatility Smile*; John Wiley & Sons: Nashville, TN, USA, 2016.
45. Das, S.R.; Sundaram, R.K. Of smiles and smirks: A term structure perspective. *J. Financ. Quant. Anal.* **1999**, *34*, 211. [CrossRef]
46. Carr, P.; Wu, L. The finite moment log stable process and option pricing. *J. Financ.* **2003**, *58*, 753–777. [CrossRef]
47. Comte, F.; Renault, E. Long memory in continuous-time stochastic volatility models. *Math. Financ.* **1998**, *8*, 291–323. [CrossRef]
48. Funahashi, H.; Kijima, M. Does the Hurst index matter for option prices under fractional volatility? *Ann. Financ.* **2017**, *13*, 55–74. [CrossRef]
49. Guyon, J.; El Amrani, M. Does the term-structure of equity at-the-money skew really follow a power law? *SSRN Electron. J.* **2022**. [CrossRef]
50. Delemotte, J.; De Marco, S.; Segonne, F. Yet another analysis of the SP500 at-the-money skew: Crossover of different power-law behaviours. *SSRN Electron. J.* **2023**. [CrossRef]
51. Fouque, J.P.; Papanicolaou, G.; Sircar, R.; Sølna, K. Maturity cycles in implied volatility. *Financ. Stochastics* **2004**, *8*. [CrossRef]
52. Gatheral, J.; Jaisson, T.; Rosenbaum, M. Volatility is rough. *Quant. Financ.* **2018**, *18*, 933–949. [CrossRef]
53. Ghysels, E.; Harvey, A.C.; Renault, E. Stochastic volatility. In *Handbook of Statistics*; Elsevier: Amsterdam, The Netherlands, 1996; pp. 119–191.
54. Frey, R. Derivative Asset Analysis in Models with Level-Dependent and Stochastic Volatility. *CWI Q.* **1997**, *10*, 1–34.
55. Cont, R. Empirical properties of asset returns: Stylized facts and statistical issues. *Quant. Financ.* **2001**, *1*, 223–236. [CrossRef]
56. Zumbach, G. *Discrete Time Series, Processes, and Applications in Finance*, 2013rd ed.; Springer: Berlin, Germany, 2012.
57. Mandelbrot, B. The Variation of Certain Speculative Prices. *J. Bus.* **1963**, *36*, 394–419. [CrossRef]
58. Fama, E.F. The Behavior of Stock-Market Prices. *J. Bus.* **1965**, *38*, 34–105. [CrossRef]
59. Cootner, P. *Random Character of Stock Market Prices*; MIT Press: London, UK, 1967.
60. Barndorff-Nielsen, O.E. Normal inverse Gaussian distributions and stochastic volatility modelling. *Scand. J. Stat. Theory Appl.* **1997**, *24*, 1–13. [CrossRef]
61. Prause, K. The Generalized Hyperbolic Model: Estimation, Financial Derivatives, and Risk Measures. Ph.D. Thesis, Albert-Ludwigs-Universit, Freiburg, Germany, 1999.
62. Bouchaud, J.P.; Potters, M. *Theory of Financial Risks: From Statistical Physics to Risk Management*; Cambridge University Press: Cambridge, UK, 2000.
63. Cont, R.; Potters, M.; Bouchaud, J.P. Scaling in stock market data: Stable laws and beyond. *SSRN Electron. J.* **1997**. [CrossRef]
64. Eom, C.; Kaizoji, T.; Scalas, E. Fat tails in financial return distributions revisited: Evidence from the Korean stock market. *Phys. A* **2019**, *526*, 121055. [CrossRef]
65. Black, F. *Studies of Stock Price Volatility Changes*; Business and Economics Section of the American Statistical Association: Washington, DC, USA, 1976; pp. 177–181.
66. Christie, A. The stochastic behavior of common stock variances: Value, leverage and interest rate effects. *J. Financ. Econ.* **1982**, *10*, 407–432. [CrossRef]
67. Cheung, Y.W.; Ng, L.K. Stock price dynamics and firm size: An empirical investigation. *J. Financ.* **1992**, *47*, 1985–1997. [CrossRef]
68. Duffee, G. Stock returns and volatility a firm-level analysis. *J. Financ. Econ.* **1995**, *37*, 399–420. [CrossRef]
69. Aït-Sahalia, Y.; Fan, J.; Li, Y. The leverage effect puzzle: Disentangling sources of bias at high frequency. *J. Financ. Econ.* **2013**, *109*, 224–249. [CrossRef]
70. Figlewski, S.; Wang, X. Is the "leverage effect" a leverage effect? *SSRN Electron. J.* **2001**. [CrossRef]
71. Hasanhodzic, J.; Lo, A.W. On Black's leverage effect in firms with no leverage. *J. Portf. Manag.* **2019**, *46*, 106–122. [CrossRef]
72. Zumbach, G. Volatility conditional on price trends. *Quant. Financ.* **2010**, *10*, 431–442. [CrossRef]

73. Fouque, J.P.; Papanicolaou, G.; Sircar, K.R. *Derivatives in Financial Markets with Stochastic Volatility*; Cambridge University Press: Cambridge, UK, 2000.

74. Cont, R. Volatility clustering in financial markets: Empirical facts and agent-based models. In *Long Memory in Economics*; Springer: Berlin/Heidelberg, Germany, 2006; pp. 289–309.

75. Bollerslev, T.; Ole Mikkelsen, H. Modeling and pricing long memory in stock market volatility. *J. Econom.* **1996**, *73*, 151–184. [CrossRef]

76. Breidt, F.J.; Crato, N.; de Lima, P. The detection and estimation of long memory in stochastic volatility. *J. Econom.* **1998**, *83*, 325–348. [CrossRef]

77. Ding, Z.; Granger, C.W.J. Modeling volatility persistence of speculative returns: A new approach. *J. Econom.* **1996**, *73*, 185–215. [CrossRef]

78. Ding, Z.; Granger, C.W.J.; Engle, R.F. A long memory property of stock market returns and a new model. *J. Empir. Financ.* **1993**, *1*, 83–106. [CrossRef]

79. Guillaume, D.M.; Dacorogna, M.M.; Davé, R.R.; Müller, U.A.; Olsen, R.B.; Pictet, O.V. From the bird's eye to the microscope: A survey of new stylized facts of the intra-daily foreign exchange markets. *Financ. Stochastics* **1997**, *1*, 95–129. [CrossRef]

80. Beran, J. *Statistics for Long-Memory Processes*; Chapman and Hall/CRC: Boca Raton, FL, USA, 1994.

81. Samorodnitsky, G. *Stochastic Processes and Long Range Dependence*; Springer International Publishing: Cham, Switzerland, 2016.

82. Mikosch, T.; Starica, C. Is it really long memory we see in financial returns? *Extrem. Integr. Risk Manag.* **2000**, *12*, 149–168.

83. Willinger, W.; Taqqu, M.S.; Teverovsky, V. Stock market prices and long-range dependence. *Financ. Stochastics* **1999**, *3*, 1–13. [CrossRef]

84. Lobato, I.N.; Velasco, C. Long memory in stock-market trading volume. *J. Bus. Econ. Stat.* **2000**, *18*, 410–427. [CrossRef]

85. Ross, S.A. A Simple Approach to the Valuation of Risky Streams. *J. Bus.* **1978**, *51*, 453–475. [CrossRef]

86. Harrison, J.M.; Kreps, D.M. Martingales and arbitrage in multiperiod securities markets. *J. Econ. Theory* **1979**, *20*, 381–408. [CrossRef]

87. Harrison, J.M.; Pliska, S.R. Martingales and stochastic integrals in the theory of continuous trading. *Stoch. Process. Their Appl.* **1981**, *11*, 215–260. [CrossRef]

88. Kreps, D.M. Arbitrage and equilibrium in economies with infinitely many commodities. *J. Math. Econ.* **1981**, *8*, 15–35. [CrossRef]

89. Delbaen, F.; Schachermayer, W. A general version of the fundamental theorem of asset pricing. *Math. Ann.* **1994**, *300*, 463–520. [CrossRef]

90. Delbaen, F.; Schachermayer, W. The fundamental theorem of asset pricing for unbounded stochastic processes. *Math. Ann.* **1998**, *312*, 215–250. [CrossRef]

91. Delbaen, F.; Schachermayer, W. *Mathematics of Arbitrage*; Springer: New York, NY, USA, 2006.

92. Schachermayer, W. The fundamental theorem of asset pricing. In *Handbook of the Fundamentals of Financial Decision Making*; World Scientific: Singapore, 2013; pp. 31–48.

93. Cox, J.C. Notes on Option pricing I: The constant elasticity of variance option pricing model. *J. Portf. Manag.* **1997**, *23*, 15–17. [CrossRef]

94. Cox, J.C.; Ross, S.A. The valuation of options for alternative stochastic processes. *J. Financ. Econ.* **1976**, *3*, 145–166. [CrossRef]

95. Emanuel, D.C.; MacBeth, J.D. Further results on the constant elasticity of variance call option pricing model. *J. Financ. Quant. Anal.* **1982**, *17*, 533. [CrossRef]

96. Linetsky, V.; Mendoza, R. Constant elasticity of variance (CEV) diffusion model. *Encycl. Quant. Financ.* **2010**. [CrossRef]

97. Yamada, T.; Watanabe, S. On the uniqueness of solutions of stochastic differential equations. *Kyoto J. Math.* **1971**, *11*, 155–167. [CrossRef]

98. Cherny, A.S.; Engelbert, H.J. *Singular Stochastic Differential Equations*; Springer: Berlin/Heidelberg, Germany, 2005.

99. Andersen, L.B.G.; Piterbarg, V.V. Moment explosions in stochastic volatility models. *Financ. Stochastics* **2006**, *11*, 29–50. [CrossRef]

100. Derman, E.; Kani, I. Riding on the Smile. *Risk* **1994**, *7*, 32–39.

101. Dupire, B. Pricing with a smile. *Risk* **1994**, *7*, 18–20.

102. Bergomi, L. *Stochastic Volatility Modeling*; Chapman & Hall/CRC: Oakville, MO, USA, 2016.

103. Buraschi, A.; Jackwerth, J. The price of a smile: Hedging and spanning in option markets. *Rev. Financ. Stud.* **2001**, *14*, 495–527. [CrossRef]

104. Dumas, B.; Fleming, J.; Whaley, R.E. Implied volatility functions: Empirical tests. *J. Financ.* **1998**, *53*, 2059–2106. [CrossRef]

105. Aït-Sahalia, Y.; Jacod, J. Testing Whether Volatility Can Be Written As a Function of the Asset Price. In *Workshop on Asset Pricing and Risk Management*; Duke University: Durham, NC, USA, 2019.

106. Hull, J.; White, A. The pricing of options on assets with stochastic volatilities. *J. Financ.* **1987**, *42*, 281–300. [CrossRef]

107. Scott, L.O. Option pricing when the variance changes randomly: Theory, estimation, and an application. *J. Financ. Quant. Anal.* **1987**, *22*, 419. [CrossRef]

108. Stein, E.M.; Stein, J.C. Stock Price Distributions with Stochastic Volatility: An Analytic Approach. *Rev. Financ. Stud.* **1991**, *4*, 727–752. [CrossRef]

109. Schöbel, R.; Zhu, J. Stochastic volatility with an Ornstein–Uhlenbeck process: An extension. *Rev. Financ.* **1999**, *3*, 23–46. [CrossRef]

110. Heston, S.L. A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options. *Rev. Financ. Stud.* **1993**, *6*, 327–343. [CrossRef]

111. Cox, J.C.; Ingersoll, J.E.; Ross, S.A. A theory of the term structure of interest rates. *Econom. J. Econom. Soc.* **1985**, *53*, 385–407. [CrossRef]
112. Benhamou, E.; Gobet, E.; Miri, M. Time dependent Heston model. *SIAM J. Financ. Math.* **2010**, *1*, 289–325. [CrossRef]
113. Goutte, S.; Ismail, A.; Pham, H. Regime-switching stochastic volatility model: Estimation and calibration to VIX options. *Appl. Math. Financ.* **2017**, *24*, 38–75. [CrossRef]
114. Melino, A.; Turnbull, S.M. Pricing foreign currency options with stochastic volatility. *J. Econom.* **1990**, *45*, 239–265. [CrossRef]
115. Melino, A.; Turnbull, S.M. Misspecification and the pricing and hedging of long-term foreign currency options. *J. Int. Money Financ.* **1995**, *14*, 373–393. [CrossRef]
116. Hagan, P.; Kumar, D.; Lesniewski, A.; Woodward, D. Managing Smile Risk. *Wilmott Mag.* **2002**, *1*, 84–108.
117. Lewis, A.L. *Option Valuation under Stochastic Volatility*; Finance Press: Vancouver, BC, Canada, 2000.
118. Carr, P.; Sun, J. A new approach for option pricing under stochastic volatility. *Rev. Deriv. Res.* **2007**, *10*, 87–150. [CrossRef]
119. Baldeaux, J.; Badran, A. Consistent Modelling of VIX and Equity Derivatives Using a 3/2 plus Jumps Model. *Appl. Math. Financ.* **2014**, *21*, 299–312. [CrossRef]
120. Javaheri, A. The volatility process: A study of stock market dynamics via para- metric stochastic volatility models and a comparison to the information embedded in option prices. Ph.D. Thesis, Ecole de Mines de Paris, Paris, France, 2004.
121. Bakshi, G.; Ju, N.; Ou-Yang, H. Estimation of continuous-time models with an application to equity volatility dynamics. *J. Financ. Econ.* **2006**, *82*, 227–249. [CrossRef]
122. Gatheral, J. Consistent Modeling of SPX and VIX options. In Proceedings of the 5th World Congress of the Bachelier Finance Society, London, UK, 15–19 July 2008.
123. Fouque, J.P.; Papanicolaou, G.; Sircar, R.; Sølna, K. Multiscale stochastic volatility asymptotics. *Multiscale Model. Simul.* **2003**, *2*, 22–42. [CrossRef]
124. Fouque, J.P.; Saporito, Y.F. Heston stochastic vol-of-vol model for joint calibration of VIX and S&P 500 options. *Quant. Financ.* **2018**, *18*, 1003–1016. [CrossRef]
125. Renault, E.; Touzi, N. Option hedging and implied volatilities in a stochastic volatility model. *Math. Financ.* **1996**, *6*, 279–302. [CrossRef]
126. Alòs, E.; León, J.A.; Vives, J. On the short-time behavior of the implied volatility for jump-diffusion models with stochastic volatility. *Financ. Stochastics* **2007**, *11*, 571–589. [CrossRef]
127. Kolmogorov, A. Wiener Spirals and some other interesting curves in Hilbert space. *Dokl Akad. Nauk SSSR* **1940**, *26*, 115–118.
128. Mandelbrot, B.B.; Van Ness, J.W. Fractional Brownian motions, fractional noises and applications. *SIAM Rev.* **1968**, *10*, 422–437. [CrossRef]
129. Mishura, Y. *Stochastic Calculus for Fractional Brownian Motion and Related Processes*, 2008th ed.; Springer: Berlin/Heidelberg, Germany, 2007.
130. Rosenbaum, M. Estimation of the volatility persistence in a discretely observed diffusion model. *Stoch. Process. Their Appl.* **2008**, *118*, 1434–1462. [CrossRef]
131. Chronopoulou, A.; Viens, F.G. Estimation and pricing under long-memory stochastic volatility. *Ann. Financ.* **2012**, *8*, 379–403. [CrossRef]
132. Bezborodov, V.; Persio, L.D.; Mishura, Y. Option Pricing with Fractional Stochastic Volatility and Discontinuous Payoff Function of Polynomial Growth. *Methodol. Comput. Appl. Probab.* **2018**, *21*, 331–366. [CrossRef]
133. Mishura, Y.; Yurchenko-Tytarenko, A. Fractional Cox–Ingersoll–Ross process with non-zero "mean". *Mod. Stochastics Theory Appl.* **2018**, *5*, 99–111. [CrossRef]
134. Mishura, Y.; Yurchenko-Tytarenko, A. Fractional Cox–Ingersoll–Ross process with small Hurst indices. *Mod. Stochastics Theory Appl.* **2018**, *6*, 13–39. [CrossRef]
135. Mishura, Y.; Yurchenko-Tytarenko, A. Approximating expected value of an option with non-Lipschitz payoff in fractional Heston-type model. *Int. J. Theor. Appl. Financ.* **2020**, *23*, 2050031. [CrossRef]
136. Bäuerle, N.; Desmettre, S. Portfolio optimization in fractional and rough Heston models. *SIAM J. Financ. Math.* **2020**, *11*, 240–273. [CrossRef]
137. Alòs, E.; Garcia Lorite, D. *Malliavin Calculus in Finance: Theory and Practice*; CRC Press: London, UK, 2021.
138. Nualart, D. *The Malliavin Calculus and Related Topics*; Springer: Berlin/Heidelberg, Germany, 2006.
139. Gihman, I.; Skorokhod, A. *The Theory of Stochastic Processes I*; Springer: Berlin/Heidelberg, Germany, 2004.
140. Fukasawa, M. Volatility has to be rough. *Quant. Financ.* **2021**, *21*, 1–8. [CrossRef]
141. Rough Volatility Literature. Available online: https://sites.google.com/site/roughvol/home/rough-volatility-literature?authuser=0 (accessed on 30 August 2023).
142. Bayer, C.; Friz, P.; Gatheral, J. Pricing under rough volatility. *Quant. Financ.* **2016**, *16*, 887–904. [CrossRef]
143. Jacquier, A.; Martini, C.; Muguruza, A. On VIX futures in the rough Bergomi model. *Quant. Financ.* **2018**, *18*, 45–61. [CrossRef]
144. Guyon, J. On the Joint Calibration of SPX and VIX Options: A Dispersion-Constrained Martingale Transport Approach. In Proceedings of the Research in Options 2019, IMPA, Rio de Janeiro, Brazil, 15–19 July 2019.
145. Lacombe, C.; Muguruza, A.; Stone, H. Asymptotics for volatility derivatives in multi-factor rough volatility models. *Math. Financ. Econ.* **2021**, *15*, 545–577. [CrossRef]
146. Fukasawa, M.; Gatheral, J. A rough SABR formula. *Front. Math. Financ.* **2022**, *1*, 81. [CrossRef]

147. Harms, P.; Stefanovits, D. Affine representations of fractional processes with applications in mathematical finance. *Stoch. Process. Their Appl.* **2019**, *129*, 1185–1228. [CrossRef]

148. Garnier, J.; Sølna, K. Optimal hedging under fast-varying stochastic volatility. *Siam J. Financ. Math.* **2020**, *11*, 274–325. [CrossRef]

149. El Euch, O.; Rosenbaum, M. The characteristic function of rough Heston models. *Math. Financ.* **2019**, *29*, 3–38. [CrossRef]

150. Gatheral, J.; Jusselin, P.; Rosenbaum, M. The quadratic rough Heston model and the joint S&P 500/VIX smile calibration problem. *SSRN Electron. J.* **2020**. [CrossRef]

151. Rosenbaum, M.; Zhang, J. Deep calibration of the quadratic rough Heston model. *arXiv* **2021**, arXiv:2107.01611.

152. Dandapani, A.; Jusselin, P.; Rosenbaum, M. From quadratic Hawkes processes to super-Heston rough volatility models with Zumbach effect. *Quant. Financ.* **2021**, *21*, 1235–1247. [CrossRef]

153. El Euch, O.; Gatheral, J.; Radoičić, R.; Rosenbaum, M. The Zumbach effect under rough Heston. *Quant. Financ.* **2020**, *20*, 235–241. [CrossRef]

154. El Euch, O.; Fukasawa, M.; Rosenbaum, M. The microstructural foundations of leverage effect and rough volatility. *Financ. Stochastics* **2018**, *22*, 241–280. [CrossRef]

155. Rogers, L.C.G. Things We think We Know (Working Paper). 2019. Available online: https://www.skokholm.co.uk/wp-content/uploads/2019/11/TWTWKpaper.pdf (accessed on 23 August 2023).

156. Cont, R.; Das, P. Rough volatility: Fact or artefact? *arXiv* **2022**, arXiv:2203.13820.

157. Fukasawa, M.; Takabatake, T.; Westphal, R. Is volatility rough? *arXiv* **2019**, arXiv:1905.04852.

158. Fukasawa, M.; Takabatake, T.; Westphal, R. Consistent estimation for fractional stochastic volatility model under high-frequency asymptotics. *Math. Financ. Int. J. Math. Stat. Financ. Econ.* **2022**, *32*, 1086–1132. [CrossRef]

159. Bolko, A.E.; Christensen, K.; Pakkanen, M.S.; Veliyev, B. A GMM approach to estimate the roughness of stochastic volatility. *J. Econom.* **2023**, *235*, 745–778. [CrossRef]

160. Funahashi, H.; Kijima, M. A solution to the time-scale fractional puzzle in the implied volatility. *Fractal Fract.* **2017**, *1*, 14. [CrossRef]

161. Bennedsen, M.; Lunde, A.; Pakkanen, M.S. Decoupling the short- and long-term behavior of stochastic volatility. *J. Financ. Econom.* **2022**, *20*, 961–1006. [CrossRef]

162. Kubilius, K.; Mishura, Y.; Ralchenko, K. *Parameter Estimation in Fractional Diffusion Models*; Springer International Publishing: Cham, Switzerland, 2017.

163. Corlay, S.; Lebovits, J.; Lévy Véhel, J. Multifractional stochastic volatility models. *Math. Financ.* **2014**, *24*, 364–402. [CrossRef]

164. Ayache, A.; Peng, Q. Stochastic volatility and multifractional Brownian motion. In *Stochastic Differential Equations and Processes*; Springer: Berlin/Heidelberg, Germany, 2012; pp. 211–237.

165. Catalini, G.; Pacchiarotti, B. Asymptotics for multifactor Volterra type stochastic volatility models. *Stoch. Anal. Appl.* **2022**, 1–31. [CrossRef]

166. Merino, R.; Pospíšil, J.; Sobotka, T.; Sottinen, T.; Vives, J. Decomposition formula for rough Volterra stochastic volatility models. *Int. J. Theor. Appl. Financ.* **2021**, *24*, 2150008. [CrossRef]

167. Di Nunno, G.; Mishura, Y.; Yurchenko-Tytarenko, A. Option pricing in Volterra sandwiched volatility model. *arXiv* **2022**, arXiv:2209.10688.

168. Di Nunno, G.; Yurchenko-Tytarenko, A. Sandwiched Volterra Volatility model: Markovian approximations and hedging. *arXiv* **2022**, arXiv:2209.13054.

169. Biagini, F.; Guasoni, P.; Pratelli, M. Mean-Variance Hedging for Stochastic Volatility Models. *Math. Financ.* **2000**, *10*, 109–123. [CrossRef]

170. Alòs, E.; Rolloos, F.; Shiraya, K. Forward start volatility swaps in rough volatility models. *arXiv* **2022**, arXiv:2207.10370.

171. Ocone, D.L.; Karatzas, I. A generalized Clark representation formula, with application to optimal portfolios. *Stochastics Stochastics Rep.* **1991**, *34*, 187–220. [CrossRef]

172. Edwards, T.; Lazzara, C. *Realized Volatility Indices: Measuring Market Risk*; Research for S&P Dow Jones Indices; McGraww Hill Financial: New York, NY, USA, 2016.

173. Abi Jaber, E.; El Euch, O. Multifactor approximation of rough volatility models. *SIAM J. Financ. Math.* **2019**, *10*, 309–349. [CrossRef]

174. Abi Jaber, E.; Illand, C.; Li, S. Joint SPX-VIX calibration with Gaussian polynomial volatility models: Deep pricing with quantization hints. *arXiv* **2022**, arXiv:2212.08297.

175. Di Nunno, G.; Mishura, Y.; Yurchenko-Tytarenko, A. Drift-implicit Euler scheme for sandwiched processes driven by Hölder noises. *Numer. Algorithms* **2023**, *93*, 459–491. [CrossRef]

176. CBOE. *Cboe Global Markets Annual Report 2022*; Cboe Global Markets, Inc.: Chicago, IL, USA, 2023.

177. Carr, P.; Lee, R. Volatility derivatives. *Annu. Rev. Financ. Econ.* **2009**, *1*, 319–339. [CrossRef]

178. Gastineau, G.L. An index of listed option premiums. *Financ. Anal. J.* **1977**, *33*, 70–75. [CrossRef]

179. Cox, J.C.; Rubinstein, M. *Options Markets*; Pearson: Upper Saddle River, NJ, USA, 1985.

180. Brenner, M.; Galai, D. New financial instruments for hedging changes in volatility. *Financ. Anal. J.* **1989**, *45*, 61–65. [CrossRef]

181. Fleming, J.; Ostdiek, B.; Whaley, R.E. Predicting stock market volatility: A new measure. *J. Futur. Mark.* **1995**, *15*, 265–302. [CrossRef]

182. Breeden, D.T.; Litzenberger, R.H. Prices of state-contingent claims implicit in option prices. *J. Bus.* **1978**, *51*, 621–651. [CrossRef]

183. Dupire, B. Arbitrage Pricing with Stochastic Volatility. 1993. Available online: https://cims.nyu.edu/~essid/ctf/stochvol.pdf (accessed on 30 August 2023).
184. Carr, P.; Madan, D. Towards a theory of volatility trading. In *Volatility*; Jarrow, R., Ed.; Risk Publications: London, UK, 1998; pp. 417–427.
185. Guyon, J. Dispersion-constrained martingale Schrödinger problems and the exact joint S&P 500/VIX smile calibration puzzle. *SSRN Electron. J.* **2021**. [CrossRef]
186. VIX Index Historical Data. Available online: https://www.cboe.com/tradable_products/vix/vix_historical_data/ (accessed on 26 September 2023).
187. CBOE. Volatility Index Methodology: Cboe Volatility Index. 2023. Available online: https://cdn.cboe.com/api/global/us_indices/governance/Volatility_Index_Methodology_Cboe_Volatility_Index.pdf (accessed on 30 August 2023).
188. Kwok, Y.K.; Zheng, W. *Pricing Models of Volatility Products and Exotic Variance Derivatives*; Chapman & Hall/CRC: Philadelphia, PA, USA, 2022.
189. Alòs, E.; García-Lorite, D.; Gonzalez, A.M. On smile properties of volatility derivatives: Understanding the VIX skew. *SIAM J. Financ. Math.* **2022**, *13*, 32–69. [CrossRef]
190. Kokholm, T.; Stisen, M. Joint pricing of VIX and SPX options with stochastic volatility and jump models. *J. Risk Financ.* **2015**, *16*, 27–48. [CrossRef]
191. Rømer, S.E. Empirical analysis of rough and classical stochastic volatility models to the SPX and VIX markets. *Quant. Financ.* **2022**, *22*, 1805–1838. [CrossRef]
192. Abi Jaber, E.; Illand, C.; Li, S. The quintic Ornstein-Uhlenbeck volatility model that jointly calibrates SPX & VIX smiles. *arXiv* **2022**, arXiv:2212.10917.
193. Cont, R.; Kokholm, T. A Consistent Pricing Model for Index Options and Volatility Derivatives. *Math. Financ.* **2013**, *23*, 248–274. [CrossRef]
194. Guyon, J. The joint S&P 500/VIX smile calibration puzzle solved. *SSRN Electron. J.* **2019**. [CrossRef]
195. Guyon, J.; Bourgey, F. Fast exact joint S&P 500/VIX smile calibration in discrete and continuous time. *SSRN Electron. J.* **2022**. [CrossRef]
196. Guyon, J.; Mustapha, S. Neural joint S&P 500/VIX smile calibration. *SSRN Electron. J.* **2022**. [CrossRef]
197. Aït-Sahalia, Y.; Karaman, M.; Mancini, L. The term structure of equity and variance risk premia. *J. Econom.* **2020**, *219*, 204–230. [CrossRef]
198. Andersen, T.G.; Bondarenko, O.; Gonzalez-Perez, M.T. Exploring return dynamics via corridor implied volatility. *Rev. Financ. Stud.* **2015**, *28*, 2902–2945. [CrossRef]
199. Carr, P.; Lee, R.; Wu, L. Variance swaps on time-changed Lévy processes. *Financ. Stochastics* **2012**, *16*, 335–355. [CrossRef]
200. Martin, I. *Simple Variance Swaps*; NBER Working Paper Series; National Bureau of Economic Research: Cambridge, MA, USA, 2011.
201. Guyon, J.; Lekeufack, J. Volatility is (mostly) path-dependent. *Quant. Financ.* **2023**, *23*, 1221–1258. [CrossRef]
202. Buehler, H.; Gonon, L.; Teichmann, J.; Wood, B. Deep hedging. *Quant. Financ.* **2019**, *19*, 1271–1291. [CrossRef]
203. Bühler, H.; Horvath, B.; Lyons, T.; Arribas, I.P.; Wood, B. A data-driven market simulator for small data environments. *arXiv* **2020**, arXiv:2006.14498.
204. Cuchiero, C.; Gazzani, G.; Möller, J.; Svaluto-Ferro, S. Joint calibration to SPX and VIX options with signature-based models. *arXiv* **2023**, arXiv:2301.13235.
205. Cuchiero, C.; Gazzani, G.; Svaluto-Ferro, S. Signature-based models: Theory and calibration. *SIAM J. Financ. Math.* **2023**, *14*, 910–957. [CrossRef]