

Implied Volatility in Strict Local Martingale Models*

Antoine Jacquier[†] and Martin Keller-Ressel[‡]

Abstract. We consider implied volatilities in asset pricing models, where the discounted underlying is a strict local martingale under the pricing measure. Our main result gives an asymptotic expansion of the right wing of the implied volatility smile and shows that the strict local martingale property can be determined from this expansion. This result complements the well-known asymptotic results of Lee and of Benaim and Friz, which apply only to true martingales. This also shows that “price bubbles” in the sense of strict local martingale behavior can in principle be detected by an analysis of implied volatility. Finally we relate our results to left-wing expansions of implied volatilities in models with mass at zero by a duality method based on an absolutely continuous measure change.

Key words. implied volatility, asymptotic methods, strict local martingale, asset price bubbles

AMS subject classifications. 91G20, 60G48

DOI. 10.1137/16M1069651

1. Introduction. Most models in mathematical finance are bound to the paradigm that discounted asset prices are true martingales under the risk neutral measure. However, no-arbitrage theory as developed by Delbaen and Schachermayer [12, 14, 15] actually allows for the more general case of discounted asset prices being local martingales, while still remaining consistent with absence of arbitrage in the sense of no free lunch with vanishing risk (NFLVR). Processes which fall in the latter category but not in the former, i.e., which are local martingales but not true martingales, are usually called strict local martingales. Although not trivial to construct, the appearance of strict local martingales in certain local and stochastic volatility models has been noted in [32, 36, 45, 13]. In contrast to true martingale models, where the current asset price equals the risk-neutral expectation of the future discounted asset price (“fundamental price”), these two values differ from each other in strict local martingale models. For this very reason, strict local martingale models have been interpreted as models for markets with price bubbles in [37, 24] or in the review paper [41]. For one-dimensional Itô diffusions without drift, a precise characterization of strict local martingales was obtained in [8, 16, 18]; this was recently used in a series of papers by Jarrow, Kchia, and Protter [28, 29, 30] to propose a statistical test based on realized volatility to determine whether a given underlying (LinkedIn’s stock and gold) exhibits a price bubble. In the context of option pricing, it has been shown that the strict local martingale property

*Received by the editors April 7, 2016; accepted for publication (in revised form) July 24, 2017; published electronically January 30, 2018.

<http://www.siam.org/journals/sifin/9-1/M106965.html>

Funding: The first author’s research was supported by EPSRC First Grant EP/M008436/1. The second author’s research was supported by the Excellence Initiative of the German Research Foundation (DFG) under grant ZUK 64.

[†]Department of Mathematics, Imperial College London, London, SW7 1AZ, UK (a.jacquier@imperial.ac.uk).

[‡]Fachrichtung Mathematik, Institut für Mathematische Stochastik, TU Dresden, 01062 Dresden, Germany (Martin.Keller-Ressel@tu-dresden.de).

leads to unexpected and counterintuitive behavior. In particular, put-call parity fails, classical no-static-arbitrage bounds are no longer valid and there is some ambiguity about the proper valuation of derivatives with unbounded payoffs, such as calls. We refer to [10] and a more elaborate discussion in section 2.2.

In this paper we focus on the properties of implied volatilities and of the resulting implied volatility surface in asset pricing models with the strict local martingale property. With the exception of [46], which discusses long-time behavior of implied volatilities, this is to our knowledge the first paper to discuss implied volatilities without imposing the true martingale assumption on the underlying. Due to the failure of put-call parity and no-static-arbitrage bounds, it turns out that put- and call-implied volatility has to be distinguished and that call-implied volatility does not always exist; see Theorem 3.1. Following an idea of Cox and Hobson [10], we introduce the “fully collateralized call,” which restores put-call parity and leads again to equality of put-implied and call-implied volatilities. Even when put-call parity is restored, our main result, Theorem 3.3, shows that the right wing of the implied volatility smile in strict local martingale models exhibits an asymptotic behavior which is fundamentally different from true martingale models. This also shows that in principle, strict local martingale models can be distinguished from true martingale models by analyzing the implied volatility surface. Moreover, this result complements the well-known results of Lee [35] and Benaim and Friz [4] on the behavior of the wings of the implied volatility smile in true (nonnegative) martingale models.

In section 4, we discuss a notion of duality between stock price models based on an absolutely continuous measure change. This method is well known in the context of strict local martingales [9, 13, 34, 40, 44] and puts positive strict local martingale models in duality with true martingale models with (positive) probability mass at zero. Recently, De Marco, Jacquier, and Hillairet [11] and Gulisashvili [22] showed that, for a true martingale with mass at zero, the left tail of the smile is fully determined (up to second order) by the probability weight of this very mass. Applying the duality method, the results on right-wing asymptotics of strict local martingales in this paper can be seen as a direct analogue of the left-wing asymptotics of true martingales with mass at zero.

2. Preliminaries.

2.1. Market models based on strict local martingales and stock price bubbles. Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{Q})$ be a probability space with a filtration satisfying the usual conditions and let $(S_t)_{t \geq 0}$ be a càdlàg nonnegative local $\mathbb{Q}$ -martingale starting at $S_0 = 1$. As a nonnegative local martingale, S is also a supermartingale. Hence $\mathbb{E}^{\mathbb{Q}}[S_t]$ exists, is bounded by 1, and is a decreasing function of $t \geq 0$. We set

$$(2.1) \quad \mathfrak{m}_t := 1 - \mathbb{E}^{\mathbb{Q}}[S_t]$$

and call $\mathfrak{m}_t$ the martingale defect (at $t \geq 0$) of S . Clearly, $t \mapsto \mathfrak{m}_t$ is increasing and takes values in $[0, 1]$, and S is a true martingale on $[0, T]$ if and only if $\mathfrak{m}_T = 0$. We will be mainly interested in the complementary case $\mathfrak{m}_T > 0$, in which S is a local martingale on $[0, T]$, but not a true martingale, i.e., a *strict local martingale*. The size of $\mathfrak{m}_T$ quantifies the difference to being a true martingale, hence the term “martingale defect.” The other boundary case $\mathfrak{m}_T = 1$ takes place if and only if $S_T = 0$, $\mathbb{Q}$ -almost surely. We exclude this degenerate case

from our analysis and work under the assumption that $\mathbf{m}_t < 1$ for all $t \in [0, T]$. We interpret S as the price of a stock and $\mathbb{Q}$ as a given pricing measure which determines the prices of derivative contracts. The historical/statistical measure will not play a role in our analysis and we do not make a priori assumptions on market completeness. The normalization of S_0 to 1 and the absence of discounting serve to simplify notation and arguments. Starting values $S_0 \neq 1$ can be accommodated by simple scaling and discounting by interpreting S as a forward price and by adjusting option strikes to forward strikes.

Recall that the market model $(S, \mathbb{Q})$ does not allow for arbitrage opportunities in the sense of NFLVR. Indeed, for a nonnegative discounted semimartingale stock price S , the first fundamental theorem of asset pricing states that the NFLVR condition holds if and only if there exists an equivalent probability measure under which S is a local martingale. We refer the interested reader to [14] or [15] for a precise definition of the NFLVR condition and details of the first fundamental theorem of asset pricing.

Concrete examples of market models where $\mathbf{m}_T > 0$, i.e., where S is a strict local $\mathbb{Q}$ -martingale, appear in both local volatility and stochastic volatility models [32, 36, 45]; see also section 5. With regard to option pricing, strict local martingale models can exhibit quite unexpected behavior: as we will discuss in more detail in section 2.2 below, put-call parity fails and put and call prices may violate the classic no-static-arbitrage bounds in such models. Moreover, there exist strikes for which the European call is strictly cheaper than its American counterpart [10].

Finally, note that from (2.1) the martingale defect $\mathbf{m}_t$ can be interpreted as the difference between the current asset price $S_0 = 1$ and the risk-neutral expectation of its future value S_t (its “fundamental value”). Following [24, 41] a nonzero difference therefore constitutes a *price bubble* in which traded values of assets diverge from fundamental values, a phenomenon that cannot appear under a true martingale assumption on S .

Example 2.1. Let X be the unique weak solution to the stochastic differential equation $dX_t = \sigma(X_t)dW_t$, with $X_0 = 1$, where σ satisfies the Engelbert–Schmidt conditions [33, Theorem 5.7], $\sigma(0) = 0$, and W is a Brownian motion. The behavior of X at zero and its martingale property can be read directly from the following integrability conditions on σ [8, 16, 18]:

- $X_t > 0$ for all $t \geq 0$ almost surely if and only if

$$(2.2) \quad \int_0^1 \frac{x}{\sigma^2(x)} dx = \infty;$$

- X is a strict local martingale if and only if

$$(2.3) \quad \int_1^\infty \frac{x}{\sigma^2(x)} dx < \infty.$$

2.2. Option pricing in strict local martingale models. Even though strict local martingale models are consistent with absence of arbitrage (in the sense of NFLVR), several apparent pathologies emerge in the context of option pricing [10, 24]. We will summarize the phenomena which are relevant in the context of implied volatility. First let us define call and put prices on the underlying S (and for fixed maturity T) by their risk-neutral expectations

$$C_S(x) := \mathbb{E}^{\mathbb{Q}}(S_T - e^x)_+ \quad \text{and} \quad P_S(x) := \mathbb{E}^{\mathbb{Q}}(e^x - S_T)_+,$$

where in view of our analysis of implied volatility we parametrize by log-strike $x = \log K$. In complete markets these prices are indeed the unique minimal superreplication prices of their payoffs (see [10]), but—as will be seen below—other sensible prices for puts and calls which are consistent with absence of arbitrage exist. It is easy to see that

$$(2.4) \quad C_S(x) - P_S(x) = 1 - e^x - \mathbf{m}_T$$

holds for all real numbers x , and hence put-call parity fails whenever $\mathbf{m}_T > 0$, i.e., exactly when S is a strict local martingale [10, Theorem 3.4]. A further pathology emerges when we consider price bounds. A direct application of Jensen's inequality yields the inequalities

$$(2.5) \quad (1 - \mathbf{m}_T - e^x)_+ \leq C_S(x) < 1 - \mathbf{m}_T,$$

$$(2.6) \quad (e^x - 1 + \mathbf{m}_T)_+ \leq P_S(x) < e^x$$

for the call and the put, valid for all $x \in \mathbb{R}$. For $\mathbf{m}_T = 0$ the process S is a true martingale and these bounds become the classic no-static-arbitrage bounds for calls and puts.¹ In the strict local martingale case, $\mathbf{m}_T > 0$, the lower bound for the call falls outside the no-static-arbitrage region. The lower bound for the put on the other hand increases with $\mathbf{m}_T$ and thus always remains within the no-static-arbitrage region. Note that for any $x \in \mathbb{R}$, both lower bounds can be attained, which follows from the example in section 5.1, where a strict local martingale with the property that $S_T = 1 - \mathbf{m}_T$, almost surely, is given. As a consequence of Theorem 3.1 below, it also follows that for each given strict local martingale S with defect $\mathbf{m}_T$, there exists $x^* \in [\log(\mathbf{m}_T), 0]$ such that the corresponding call price $C_S(x)$ violates the no-static-arbitrage lower bound $(1 - e^x)_+$ for all $x < x^*$.

These pathologies appear to contradict the fact that strict local martingale models are consistent with absence of arbitrage (in the NFLVR sense). Consider, for example, a call option that is valued at the lower bound in (2.5), that is, $C_S(x) = (1 - \mathbf{m}_T - e^x)_+$. Then, choosing a log-strike x such that $x \leq \log(1 - \mathbf{m}_T)$, it is possible to form a costless portfolio consisting of a long position in the call $C_S(x)$, a short position of one unit of the stock, and $\mathbf{m}_T + e^x$ in the bank account. The payoff at maturity of this portfolio is $\mathbf{m}_T + (e^x - S_T)_+ > 0$, and we have apparently constructed an arbitrage. The resolution of this paradox is that due to the short position in S the value process of the portfolio is unbounded from below and therefore not an admissible strategy in the sense of [12, Definition 2.7]. Indeed, if the portfolio value were bounded from below, then S would be bounded from above and hence a true martingale, in contradiction to the strict martingale property of S .

These considerations regarding admissibility of strategies are not entirely academic. Short positions usually require deposit of collateral, which restricts the scope of implementable hedging strategies. For this reason it has been remarked by several authors [10, 24, 38] that defining the call price by the risk-neutral expectation of its payoff is not the only economically sensible choice. We follow here the definition in [10], where the authors consider a short position in calls and argue that such positions are usually subject to collateral requirements.

¹No-static-arbitrage refers to elementary static replication arguments that do not take into account admissibility of trading strategies. This notion has to be distinguished from no-arbitrage in the sense of NFLVR, which does take into account admissibility of strategies.

Thus, the value process V of a hedging portfolio must not only replicate the call payoff at maturity T but also satisfy the collateral requirement $V_t \geq G(S_t)$ at intermediate times. Here, the function G is used to describe the amount of collateral needed in relation to the stock price. The following is proved in [10].

Theorem 2.2 (Theorem 5.2 in [10]). *Let G be a positive convex function satisfying $\limsup_{s \uparrow \infty} \frac{G(s)}{s} = \alpha$, and let H be an arbitrary payoff satisfying $H \geq G$; then the fair price (at inception) of a European option with collateral requirement described by G and with payoff $H(S_T)$ is equal to $\mathbb{E}^{\mathbb{Q}}(H(S_T)) + \alpha \mathfrak{m}_T$.*

By “fair price,” Cox and Hobson mean the smallest initial fortune required to construct a self-financing wealth process superreplicating both the payoff at maturity and the collateral requirement through the life of the contract. In the case of a European call option (with maturity T and strike e^x), $H(S_T) = (S_T - e^x)_+$, Theorem 2.2 implies that the fair price of the call under collateralization reads

$$(2.7) \quad C_S^\alpha(x) := C_S(x) + \alpha \mathfrak{m}_T.$$

We shall refer to $C_S^\alpha(x)$ as the value of an α -collateralized call. Note that only values $\alpha \in [0, 1]$ make sense due to the requirement on G . It is clear that the collateral requirement does not affect the call price in a true martingale model where $\mathfrak{m}_T = 0$. In a strict local martingale model prices of collateralized calls differ from prices of uncollateralized calls. Of particular interest is the fully collateralized call $C_S^1(x)$, which coincides with the European call price “ C^{strict} ” proposed by Madan and Yor in [38, equation (6)ff] for strict local martingale models and with the “generalized fair value” of the call discussed in [36, Chapter 5]. Also the call prices G^2 and G^1 discussed in [24] in the context of strict local martingale models correspond precisely to the uncollateralized and the fully collateralized call prices $C_S = C_S^0$ and C_S^1 , respectively. Inserting into (2.4) and (2.5) we see that

$$(2.8) \quad C_S^1(x) - P_S(x) = 1 - e^x \quad \text{and} \quad \min(\mathfrak{m}_T, 1 - e^x) \leq C_S^1(x) \leq 1,$$

that is, for the fully collateralized call price, put-call parity is restored and the pricing bounds are always within the no-static-arbitrage region.

3. Implied volatility for strict local martingales. For each $x \in \mathbb{R}$, the implied volatility for a given call price $C(x)$ is defined as the unique nonnegative solution to the equation $C_{\text{BS}}(x, \sigma) = C(x)$, where C_{BS} represents the Black–Scholes European call price with maturity T , strike e^x , and volatility σ :

$$C_{\text{BS}}(x, \sigma) := \mathcal{N}(d_+(x, \sigma)) - e^x \mathcal{N}(d_-(x, \sigma)), \quad \text{where } d_{\pm}(x, \sigma) := \frac{-x}{\sigma\sqrt{T}} \pm \frac{1}{2}\sigma\sqrt{T},$$

with $\mathcal{N}$ standing for the Gaussian cumulative distribution function. It is known that the implied volatility is a well-defined real number in $[0, \infty)$ if and only if $C(x)$ lies within the no-static-arbitrage bounds (given by the bounds (2.5) for $\mathfrak{m}_T = 0$) and that in this case it is unique. If put-call parity holds, then the definition using European put options is equivalent to that using call options. Hence in true martingale models, the implied volatility is always uniquely defined and there is no distinction between call- and put-implied volatility. The behavior of the implied volatility in such models is by now fairly well understood [20]. However,

in the strict local martingale case, surprisingly few results exist (apart from [46], which studies the large-time behavior). It turns out that in the strict local martingale setting, even existence of implied volatilities is not certain and one cannot equivalently consider call and put options. In this section, we endeavor to fill this gap by providing results on existence and uniqueness and on the asymptotic behavior for large strikes of implied volatilities in the class of strict local martingale models.

3.1. Put- and call-implied volatility. As discussed above, put-call parity fails in the strict local martingale setting (unless calls are fully collateralized), and hence call-implied volatilities have to be distinguished from put-implied volatilities. We denote by $I_S^p(x)$ the implied volatility corresponding to the price $P_S(x)$ of a put with log-strike x , written on a local martingale S , and denote by $I_S^\alpha(x)$ the implied volatility corresponding to the α -collateralized call $C_S^\alpha(x)$, namely, for each $x \in \mathbb{R}$, the unique nonnegative solution (whenever it exists) to the equation $C_{BS}(x, I_S^\alpha(x)) = C_S^\alpha(x)$. We start with the following result discussing the existence of I_S^p and I_S^α .

Theorem 3.1. *Let S be a nonnegative strict local martingale.*

- (i) *The implied volatility I_S^p of the put P_S is well defined on the whole real line.*
- (ii) *The implied volatility I_S^1 of the fully collateralized call C_S^1 is well defined on $\mathbb{R}$ and coincides with the put-implied volatility: $I_S^1(x) = I_S^p(x)$ for all $x \in \mathbb{R}$.*
- (iii) *For $\alpha \in [0, 1)$ there exists $x^*(\alpha) \leq 0$ such that the implied volatility I_S^α of the α -collateralized call is well defined on $[x^*(\alpha), +\infty)$ but not on $(-\infty, x^*(\alpha))$. The function $x_*(\alpha)$ is strictly decreasing and satisfies*

$$(3.1) \quad \log((1 - \alpha)m_T) < x^*(\alpha) \leq \log(1 - \alpha m_T).$$

For every $x \in \mathbb{R}$ the function $\alpha \mapsto I_S^\alpha(x)$ is strictly increasing on the interval where it is defined, and $I_S^\alpha(x) < I_S^p(x)$ holds for all $\alpha \in [0, 1)$ and $x \in [x^*(\alpha), +\infty)$.

Remark 3.2. Reparameterizing by log-strike and setting $K^*(\alpha) = \exp(x^*(\alpha))$ the bounds in (3.1) simplify to $(1 - \alpha)m_T < K^*(\alpha) \leq (1 - \alpha m_T)$. Even without specifying a concrete model for S , the region $\mathcal{D} := \{(\alpha, x) : \alpha \in [0, 1), x \in \mathbb{R}\}$ can be written as the disjoint union $\mathcal{D} = \mathcal{D}^A \cup \mathcal{D}^N \cup \mathcal{D}^M$, where

$$\mathcal{D}^A := \{(\alpha, x) \in \mathcal{D} : x > \log(1 - \alpha m_T)\},$$

$$\mathcal{D}^N := \{(\alpha, x) \in \mathcal{D} : x \leq \log((1 - \alpha)m_T)\},$$

$$\mathcal{D}^M := \{(\alpha, x) \in \mathcal{D} : \log((1 - \alpha)m_T) < x \leq \log(1 - \alpha m_T)\},$$

such that the implied volatility $I_S^\alpha(x)$ is always defined in $\mathcal{D}^A$, never in $\mathcal{D}^N$ and may or may not be in $\mathcal{D}^M$. It will also become clear from the proof that the region where $I_S^\alpha(x)$ is not defined is precisely the region where the call price $C_S^\alpha(x)$ violates the lower no-static-arbitrage bound.

Proof. The price of the put and the price of the fully collateralized call are always inside the no-static-arbitrage region by (2.5) and (2.6). Hence, the corresponding implied volatilities $I_S^p(x)$ and $I_S^1(x)$ are well defined for all $x \in \mathbb{R}$. The put-call parity (2.8) for the fully collateralized call and for the Black–Scholes price yields

$$P_S(x) = C_S^1(x) + e^x - 1 = C_{BS}(x, I_S^1(x)) + e^x - 1 = P_{BS}(x, I_S^1(x)).$$

Since $P_S(x) = P_{BS}(x, I_S^p(x))$, uniqueness of implied volatility implies that $I_S^1(x) = I_S^p(x)$ for all $x \in \mathbb{R}$, and claims (i) and (ii) follow.

Let now $\alpha \in [0, 1)$. The implied volatility $I^\alpha(x)$ for the call $C_S^\alpha(x)$ exists if and only if $C_S^\alpha(x)$ is inside the no-static-arbitrage region $[(1 - e^x)_+, 1)$. From (2.5) we derive the bounds

$$(3.2) \quad (1 - m_T - e^x)_+ + \alpha m_T \leq C_S^\alpha(x) < 1 + (\alpha - 1)m_T.$$

Thus, $C_S^\alpha(x)$ is in the no-static-arbitrage region if and only if $C_S^\alpha(x) \geq (1 - e^x)_+$ or equivalently if $F^\alpha(x) := C_S^\alpha(x) - (1 - e^x)_+ \geq 0$. To find the zeros of $F^\alpha(x)$ it suffices to consider $x < 0$ since $F^\alpha(x) > 0$ for all $x \geq 0$. Rewriting F^α for $x \leq 0$ as

$$F^\alpha(x) = \mathbb{E}^\mathbb{Q}(S_T - e^x)_+ + \alpha m_T - (1 - e^x) = \mathbb{E}^\mathbb{Q}(\max(S_T, e^x)) + \alpha m_T - 1,$$

we see that F^α is continuous and increasing on $(-\infty, 0]$ with $\lim_{x \downarrow -\infty} F^\alpha(x) = m_T(\alpha - 1)$. Setting $x^*(\alpha) := \inf\{x \leq 0 : F^\alpha(x) \geq 0\}$, it follows that implied volatility exists on $[x^*(\alpha), \infty)$ but not on $(-\infty, x^*(\alpha))$. It is also clear that for fixed x the function $\alpha \mapsto F^\alpha(x)$ is strictly increasing and hence that $x^*(\alpha)$ must be strictly decreasing. Moreover, considering the left limit of $F^\alpha(x)$ at $-\infty$ it follows that $x^*(\alpha) = -\infty$ for $\alpha = 1$ and $x^*(\alpha) > -\infty$ for all other $\alpha \in [0, 1)$. In the latter case it holds that $C_S^\alpha(x^*(\alpha)) = 1 - e^{x^*(\alpha)}$. Plugging the right-hand side into the bounds (3.2) and rearranging we obtain (3.1). ■

3.2. Asymptotic behavior of the implied volatility. For large strikes, the following result provides the asymptotic behavior of put- and call-implied volatilities.

Theorem 3.3. *Let S be a nonnegative strict local martingale with martingale defect $m_T \in (0, 1)$ and suppose that $\alpha > 0$. Then, as x tends to infinity, the following expansion holds for the implied volatility of the α -collateralized call:*

$$I_S^\alpha(x) = \sqrt{\frac{2x}{T}} + \frac{n_{\alpha,T}}{\sqrt{T}} + \varepsilon(x),$$

where $n_{\alpha,T} = \mathcal{N}^{-1}(\alpha m_T)$ and

$$\varepsilon(x) = \frac{\frac{1}{2}n_{\alpha,T}^2 + 1}{\sqrt{2Tx}} + \sqrt{\frac{2\pi}{T}} e^{\frac{1}{2}n_{\alpha,T}^2} C_S(x) + o(1) \left(C_S(x) + \frac{1}{\sqrt{x}} \right)$$

with $C_S(x) = \mathbb{E}^\mathbb{Q}(S_T - e^x)_+$. Since $I_S^p(x) = I_S^1(x)$, the expansion for the put-implied volatility is equal to the expansion of the fully collateralized call and obtained by setting $\alpha = 1$.

Remark 3.4. Obviously, $\varepsilon(x) \rightarrow 0$ as $x \rightarrow \infty$. However, under mild additional assumptions on the distribution of S_T much stronger results can be obtained. Assume, for example, that $\mathbb{E}^\mathbb{Q}(S_T^p) < \infty$ for some $p > 1$. Then, with $1/p + 1/q = 1$, we obtain the bound

$$C_S(x) \leq \mathbb{E}^\mathbb{Q}(S_T \mathbf{1}_{S_T \geq e^x}) \leq \|S_T\|_p \mathbb{Q}(S_T \geq e^x)^{1/q} \leq (1 - m_T)^{1/q} \|S_T\|_p e^{-x/q}$$

by Hölder's and Markov's inequality.

The above theorem should be contrasted with the following result on the implied volatility in true martingale models and for uncollateralized calls.

Corollary 3.5. *If $\alpha = 0$, then*

$$\lim_{x \uparrow \infty} \left(I_S^0(x) - \sqrt{\frac{2x}{T}} \right) = -\infty.$$

If $\mathbf{m}_T = 0$, then, for all $\alpha \in [0, 1]$,

$$\lim_{x \uparrow \infty} \left(I_S^\alpha(x) - \sqrt{\frac{2x}{T}} \right) = \lim_{x \uparrow \infty} \left(I_S^\alpha(x) - \sqrt{\frac{2x}{T}} \right) = -\infty.$$

Remark 3.6. Together, Theorem 3.3 and Corollary 3.5 show that there is a sharp distinction between the behavior of the implied volatility for large strikes in strict local martingale models and that in true martingale models. This observation can be used to detect the strict local martingale property from observed implied volatilities; see section 3.4 below.

Proof of Theorem 3.3. We consider the call-implied volatility $I_S^\alpha(x)$, which by Theorem 3.1 is well defined at least for all $x \in (0, \infty)$, for any $\alpha \in [0, 1]$. By definition, $I_S^\alpha(x)$ satisfies

$$\mathcal{N}\left(d_+(x, I_S^\alpha(x))\right) - e^x \mathcal{N}\left(d_-(x, I_S^\alpha(x))\right) = C_S^\alpha(x) = C_S(x) + \alpha \mathbf{m}_T.$$

Writing $\psi(x) = \sqrt{4\pi x} e^x \mathcal{N}(d_-(x, I_S^\alpha(x)))$ and rearranging the above equation yields

$$(3.3) \quad \mathcal{N}\left(d_+(x, I_S^\alpha(x))\right) = \alpha \mathbf{m}_T + C_S(x) + \frac{\psi(x)}{\sqrt{4\pi x}}.$$

By the arithmetic-geometric-mean inequality we have $d_-(x, \sigma) \leq -\sqrt{2x}$, and hence ψ satisfies the inequality

$$\psi(x) = \sqrt{4\pi x} e^x \mathcal{N}\left(d_-(x, I_S^\alpha(x))\right) \leq \sqrt{2\pi} e^x \phi\left(d_-(x, I_S^\alpha(x))\right) = \exp\left(-\frac{1}{2}d_+(x, I_S^\alpha(x))^2\right),$$

where ϕ denotes the standard Gaussian density and we have used the classical bound $\mathcal{N}(-x)/\phi(x) \leq x^{-1}$ on Mills ratio. Applying $\mathcal{N}^{-1}$ to both sides of (3.3) we obtain

$$(3.4) \quad d_+(x, I_S^\alpha(x)) = \mathcal{N}^{-1}\left(\alpha \mathbf{m}_T + C_S(x) + \frac{\psi(x)}{\sqrt{4\pi x}}\right).$$

The equation $d_+(x, \sigma) = u$ for $u \in \mathbb{R}$ can easily be inverted to yield

$$\sigma\sqrt{T} = u + \sqrt{u^2 + 2x} = u + \sqrt{2x} + \frac{u^2 + o(1)}{2\sqrt{2x}}$$

as x tends to zero (with u bounded). Inserting the right-hand side of (3.4) for u , we get $u^2 = n_{\alpha, T}^2 + o(1)$ and thus

$$(3.5) \quad I_S^\alpha(x)\sqrt{T} = \mathcal{N}^{-1}\left(\alpha \mathbf{m}_T + C_S(x) + \frac{\psi(x)}{\sqrt{4\pi x}}\right) + \sqrt{2x} + \frac{n_{\alpha, T}^2 + o(1)}{2\sqrt{2x}}.$$

By Taylor expansion, and noting that $\alpha \mathfrak{m}_T \in (0, 1)$, we have

$$\mathcal{N}^{-1}(\alpha \mathfrak{m}_T + \delta) = n_{\alpha, T} + \frac{1 + o(1)}{\phi(n_{\alpha, T})} \delta = n_{\alpha, T} + \left(\sqrt{2\pi} e^{\frac{1}{2} n_{\alpha, T}^2} + o(1) \right) \delta$$

as δ tends to zero. Plugging this into (3.5) gives

$$(3.6) \quad I_S^\alpha(x) = \sqrt{\frac{2x}{T}} + \frac{n_{\alpha, T}}{\sqrt{T}} + \varepsilon(x)$$

with

$$\varepsilon(x) = \frac{n_{\alpha, T}^2}{2\sqrt{2Tx}} + \frac{e^{\frac{1}{2} n_{\alpha, T}^2} \psi(x)}{\sqrt{2Tx}} + \sqrt{\frac{2\pi}{T}} e^{\frac{1}{2} n_{\alpha, T}^2} C_S(x) + o(1) \left(C_S(x) + \frac{1}{\sqrt{x}} \right),$$

and it remains to show that $\lim_{x \uparrow \infty} \psi(x) = e^{-\frac{1}{2} n_{\alpha, T}^2}$. We have already shown that (3.6) holds with $\varepsilon(x) \rightarrow 0$, so we can write

$$\psi(x) = \sqrt{4\pi x} e^x \mathcal{N} \left(-\frac{x}{V(x)} - \frac{V(x)}{2} \right) \quad \text{with} \quad V(x) = I_S^\alpha(x) \sqrt{T} = \sqrt{2x} + n_{\alpha, T} + o(1).$$

The asymptotic equivalence $\mathcal{N}(-z) \sim \phi(z)/z$ as $z \uparrow \infty$ implies that

$$\psi(x) \sim \exp \left(-\frac{1}{2} \left(\frac{x}{V(x)} + \frac{V(x)}{2} \right)^2 + x \right) = \exp \left(-\frac{1}{2} \left(\frac{x}{V(x)} - \frac{V(x)}{2} \right)^2 \right) \sim e^{-\frac{1}{2} n_{\alpha, T}^2} \quad \text{as } x \uparrow \infty,$$

i.e., $\lim_{x \uparrow \infty} \psi(x) = e^{-\frac{1}{2} n_{\alpha, T}^2}$, completing the proof. ■

Proof of Corollary 3.5. In contrast to Theorem 3.3 we now have $\alpha \mathfrak{m}_T = 0$. Even in this case, the derivation of (3.4) in the proof of Theorem 3.3 remains valid. However, the right-hand side of (3.4) now tends to $-\infty$ as $x \rightarrow \infty$. Note that the inequality $d_+(x, \sigma) \leq -u$ for some $u \geq 0$ implies that $\sigma \sqrt{T} \leq -u + \sqrt{u^2 + 2x}$. Hence,

$$\lim_{x \uparrow \infty} \left(I_S^\alpha(x) \sqrt{T} - \sqrt{2x} \right) \leq \lim_{x \uparrow \infty} \left(-u + \sqrt{2x + u^2} - \sqrt{2x} \right) = -u$$

for arbitrary $u \geq 0$, and Corollary 3.5 follows. ■

Finally, we give a nonasymptotic result on at-the-money implied volatility. The result is “dual,” in a sense that is made precise in section 4.1, to [11, Proposition 2.1], whose proof we adapt to our setting.

Proposition 3.7. *Let S be a nonnegative strict local martingale with martingale defect $\mathfrak{m}_T > 0$ and let $\alpha \in [0, 1]$. Then $I_S^\alpha(0)$, the at-the-money implied volatility of the α -collateralized call, satisfies*

$$\alpha \mathfrak{m}_t \leq 2\mathcal{N} \left(\frac{I_S^\alpha(0) \sqrt{T}}{2} \right) - 1.$$

In other words, the at-the-money implied volatility provides a model-free upper bound on the size of the martingale defect $\mathfrak{m}_T$.

Proof. In the at-the-money case $x=0$, the Black–Scholes call price degenerates to $C_{BS}(0, \sigma) = 2\mathcal{N}(\sigma \sqrt{T}/2) - 1$. On the other hand, by (2.7), we have the lower bound $C_S^\alpha(0) = \mathbb{E}^\mathbb{Q}(S_T - 1)_+ + \alpha \mathfrak{m}_T \geq \alpha \mathfrak{m}_T$, from which the proposition follows. ■

3.3. Relation to Lee's [35] and Benaim–Friz [4] asymptotics. Roger Lee [35] pioneered the analysis of the tail behavior of the implied volatility in the true martingale framework. He proved that for a given strictly positive $\mathbb{Q}$ -martingale X , the large-strike behavior of the implied volatility I_X is given by

$$(3.7) \quad \limsup_{x \uparrow \infty} \frac{I_X(x)^2 T}{x} = \psi(p^*) \in [0, 2],$$

where $p^* := \sup\{p \geq 0 : \mathbb{E}^{\mathbb{Q}}(X_T^{p+1}) < \infty\}$ and $\psi(p) \equiv 2 - 4(\sqrt{p(p+1)} - p)$. Similar results also hold for the small-strike behavior using the negative moments of X_T . Subsequently, Benaim and Friz [4] refined (3.7) by providing sufficient conditions under which the limsup can be strengthened into a genuine limit. Surprisingly, the martingale assumption is not explicitly mentioned in [4], though it is in the review paper [5]. An immediate consequence of (3.7) (or its sharpened version in [4]) is that when X is a $\mathbb{Q}$ -martingale, the large-strike slope of the total implied variance $I_X^2(x)T/x$ is equal to two if and only if $p^* = 0$, namely, if no moment strictly greater than one exists, so that the distribution of X_T has a fat right tail.

In the strict local martingale framework this one-to-one correspondence between tail weight and slope of implied volatility breaks down. Consider, for instance, the strict local martingale $dS_t = S_t^2 dW_t^{\mathbb{Q}}$, starting at $S_0 = 1$, whose density is given in (5.1). A simple Taylor expansion shows that, as s tends to infinity, the following asymptotic behavior holds for any $p \geq 0$:

$$s^p \mathbb{Q}(S_T \in ds) = \sqrt{\frac{2}{\pi T^3}} e^{-1/(2T)} s^{p-4} \{1 + \mathcal{O}(s^{-2})\} ds,$$

such that $p_S^* := \sup\{p \geq 0 : \mathbb{E}^{\mathbb{Q}}(S_T^{p+1}) < \infty\} = 2$. Should Lee's (or the Benaim–Friz) formula hold, then $\limsup_{x \uparrow \infty} I_s^\alpha(x)^2 T/x = \psi(p_S^*) < 2$ with $\alpha \in (0, 1]$. This stands in contradiction to Theorem 3.3, which states that $\limsup_{x \uparrow \infty} \frac{I_s^\alpha(x)^2 T}{x} = 2$. This example shows that the results of Lee [35] and Benaim and Friz [4] cannot hold for strict local martingales.

3.4. Testable implications of price bubbles and ε -close implied volatility. As discussed at the beginning of this paper, strict local martingale models have been advocated as models for stock price bubbles. In this regard, it is of interest to be able to test empirical data for the appearance of such bubbles. In the context of continuous Markov diffusions, such tests have been proposed and implemented by [28] (see also [26] for a similar idea) based on statistical estimation of historical volatility. Once the diffusion coefficient of the stock price is estimated and extrapolated to the whole real half-line, the integral criterion discussed in Example 2.1 is used by [28] to decide whether the underlying is a strict local martingale. Complementary to the statistical approach of [28], our results suggest different ways to test for the appearance of a stock price bubble based on implied (as opposed to historical) volatility. First, observe that in the case of nonfully-collateralized calls ($\alpha < 1$) there are simple criteria to distinguish between true and strictly local martingales, based on implied volatility. From the results presented above, it follows that unless calls are fully collateralized there is equivalence in the following statements:

- The stock price process S is a strict local martingale under the pricing measure $\mathbb{Q}$.
- Put- and call-implied volatilities are different.

- The call-implied volatility does not exist for sufficiently small strikes.

However, for fully collateralized calls ($\alpha = 1$) these criteria fail. Instead, we derive from Theorem 3.3 the following criterion:

- The implied volatility satisfies $I_S(x) = \sqrt{\frac{2x}{T}} + \frac{\mathfrak{m}_T}{\sqrt{T}} + o(1)$ for some $\mathfrak{m}_T \in \mathbb{R}$.

Note that the last criterion is necessary and sufficient for the strict local martingale property. The necessary part comes from Theorem 3.3, and it is sufficient by Corollary 3.5 applied to the case $\mathfrak{m}_T = 0$. In addition, the martingale defect $\mathfrak{m}_T$ can be extracted by setting $\mathfrak{m}_T = \mathcal{N}(\mathfrak{n}_T)$. The drawback of this criterion is that it is an *asymptotic* test, valid only for large x . This drawback is shared with the statistical test of [28], which also requires asymptotic extrapolation of the estimated diffusion coefficient. This property limits the value of the test in practical applications, since implied volatility (or call prices) can only be observed at a finite number of strikes. However, a simple argument shows that any test to determine the strict local martingale property from put-implied (or fully collateralized call-implied) volatilities is necessarily an asymptotic test in the sense that it uses arbitrarily large strike values as input: let S be a nonnegative local $\mathbb{Q}$ -martingale with localizing sequence $(\tau_n)_{n \in \mathbb{N}}$, so that the stopped processes $S_t^n := S_{t \wedge \tau_n}$ are true $\mathbb{Q}$ -martingales. The difference between put prices on S and S^n can be estimated uniformly on time-strike rectangles $\mathcal{R}_{T, \tilde{x}} := [0, T] \times (-\infty, \tilde{x})$, for any $\tilde{x} \in \mathbb{R}$, by

$$\sup_{(t,x) \in \mathcal{R}_{T, \tilde{x}}} |P_S(x, t) - P_{S^n}(x, t)| = \sup_{(t,x) \in \mathcal{R}_{T, \tilde{x}}} |\mathbb{E}^{\mathbb{Q}}(e^x - S_t)_+ - \mathbb{E}^{\mathbb{Q}}(e^x - S_t^n)_+| \leq e^{\tilde{x}} \mathbb{Q}(\tau_n \leq T).$$

The bound can be made arbitrarily small by choosing n large enough. Since put-implied volatilities depend continuously on the put price this shows that for any nonnegative strict local martingale S , time-strike rectangle $\mathcal{R}_{T, \tilde{x}}$, and $\varepsilon > 0$ we can find a true martingale model S^ε such that put-implied volatilities (as well as those implied by fully collateralized calls) are ε -close, uniformly on $\mathcal{R}_{T, \tilde{x}}$.

3.5. Large-time behavior. In [46], Tehranchi studied the large-time behavior of the total implied variance $I_S(x, T)^2 T$, where we now emphasize the dependence on the maturity T . His definition of the implied volatility, valid when the underlying is a true martingale, is, however, not fully adequate for strict local martingales. We shall therefore understand it as the put-implied volatility I_S^p or, equivalently, from Theorem 3.1, as the fully collateralized call implied volatility. His main result reads as follows.

Theorem 3.8 (Theorem 3.1 in [46]). *Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{Q})$ be a given probability space and S a nonnegative $\mathbb{Q}$ -local martingale starting at one and such that $\mathbb{Q}(S_t > 0) > 0$ for all $t \geq 0$. If S_t converges almost surely to zero as t tends to infinity, then the following holds for any real number x :*

$$I_S^p(x)^2 T = -8 \log \mathbb{E}^{\mathbb{Q}}(S_T \wedge e^x) - 4 \log \left[-\log \mathbb{E}^{\mathbb{Q}}(S_T \wedge e^x) \right] + 4x - 4 \log(\pi) + \varepsilon(x, T),$$

as T tends to infinity, where the error $\varepsilon(\cdot, T)$ vanishes as $T \uparrow \infty$ and satisfies some uniform bounds on compacts as T increases.

The assumption that S_t converges almost surely to zero is equivalent to put option prices converging to the strike e^x (by dominated convergence)—or the fully collateralized call price

converging to unity. In [42, Lemma 3.3], Rogers and Tehranchi proved that if S is a $\mathbb{Q}$ -martingale, this is also equivalent to $I_S(x, T)^2 T$ converging to infinity for any $x \in \mathbb{R}$. An immediate computation shows that this still holds if S is a (nonnegative) strict local $\mathbb{Q}$ -martingale.

Let now S be a nonnegative $\mathbb{Q}$ -supermartingale. Since S is bounded in $L^1(\mathbb{Q})$, the martingale convergence theorem [43, Theorem 69.1] ensures that the limit $S_\infty := \lim_{t \uparrow \infty} S_t$ exists almost surely in $[0, \infty)$. It is, however, not immediately clear whether $S_\infty = 0$ almost surely for a general nonnegative strict local $\mathbb{Q}$ -martingale, as required in Theorem 3.8. In the particular case where S is an Itô diffusion, Proposition 5.22 in [33] guarantees that $S_\infty = 0$, $\mathbb{Q}$ -almost surely, and therefore Theorem 3.8 holds. In a stochastic volatility context, precise conditions on the coefficients of the two-dimensional stochastic differential equation ensuring that $S_\infty = 0$ almost surely have been derived by Hobson [25].

4. Duality methods.

4.1. Duality with respect to true martingales with mass at zero. The results from the previous section can be linked in a systematic manner to the results of [11, 22] on martingales with mass at zero by exploiting the duality relation introduced below.

Definition 4.1. Let $\mathbb{Q}$ and $\mathbb{P}$ be probability measures on a filtered measure space and let $T > 0$ be a fixed time horizon. Let S be a strictly positive càdlàg local $\mathbb{Q}$ -martingale and M be a nonnegative true $\mathbb{P}$ -martingale on $[0, T]$. Denote by $\tau := \inf\{t > 0 : M_t = 0\}$ the first hitting time (of M) of zero and assume that τ is predictable and $\tau > 0$, $\mathbb{P}$ -a.s. We say that the pair $(S, \mathbb{Q})$ is in duality to $(M, \mathbb{P})$ if $\mathbb{Q}$ is absolutely continuous with respect to $\mathbb{P}$ on $\mathcal{F}_T$, with

$$\left. \frac{d\mathbb{Q}}{d\mathbb{P}} \right|_{\mathcal{F}_T} = M_T \quad \text{and} \quad S_t = \frac{1}{M_t} \quad \mathbb{P}\text{-a.s. on } \{t < \tau \wedge T\}.$$

Note that the above definition requires that S is a strictly positive local martingale, a slightly stronger assumption than the nonnegativity assumption made in section 2.1. In financial modeling, $\mathbb{Q}$ can be interpreted as the “share measure” corresponding to the stock price M under $\mathbb{P}$ or—in the context of currency models—as the “foreign measure” corresponding to the domestic measure $\mathbb{P}$ and the exchange rate process M [7, Chapter 17]. Also note that since M is a nonnegative true $\mathbb{P}$ -martingale, zero is necessarily absorbing [27, Chapter III, Lemma 3.6], and hence $M_t = 0$ for all $t \geq \tau$. For models in duality,

$$(4.1) \quad m_t = 1 - \mathbb{E}^{\mathbb{Q}}(S_t) = 1 - \mathbb{E}^{\mathbb{P}}(\mathbf{1}_{\{t < \tau\}}) = \mathbb{P}(\tau \leq t) = \mathbb{P}(M_t = 0),$$

that is, the martingale defect of S (under $\mathbb{Q}$) equals the mass at zero of M (under $\mathbb{P}$).

The following result is deep, but by now well understood, and is proved in [34, section 1] or [40, 44].

Lemma 4.2. Let $(\mathcal{F}_t)_{t \geq 0}$ be the right-continuous augmentation of a standard system and let $T > 0$ be a fixed time horizon. For any pair $(S, \mathbb{Q})$ satisfying the assumptions of Definition 4.1 there exists a dual pair $(M, \mathbb{P})$. Conversely, for any pair $(M, \mathbb{P})$ and associated stopping time τ satisfying the assumptions of Definition 4.1, there exists a dual pair $(S, \mathbb{Q})$.

Let us remark that the difficult direction is going from $(S, \mathbb{Q})$ to $(M, \mathbb{P})$ in the case where S is a strict local martingale; the required construction of $\mathbb{P}$ relies on the Föllmer exit measure, first introduced in [19] and mentioned in [39]. Going from $(M, \mathbb{P})$ to $(S, \mathbb{Q})$ is easier, and an early proof under the assumption of continuity can be found in Delbaen and Schachermayer [13]. For the technical condition of $(\mathcal{F}_t)_{t \geq 0}$ being the right continuous augmentation of a standard system we refer to [34, Lemma 1.5], where it is shown that the natural filtration of the coordinate process on a “modified Skorokhod space” $D'(\mathbb{R}_{\geq 0}, \mathbb{R} \cup \{+\infty\})$ constitutes a standard system that is suitable for most applications.

For models in duality we have the following simple characterization of the strict local martingale property of S under $\mathbb{Q}$.

Lemma 4.3. *Let $(M, \mathbb{P})$ and $(S, \mathbb{Q})$ be market models in duality with time horizon $T > 0$. The following are equivalent:*

- (i) S is a strict local $\mathbb{Q}$ -martingale on $[0, T]$;
- (ii) M_T has mass at zero, i.e., $\mathbb{P}(M_T = 0) > 0$;
- (iii) $\mathfrak{m}_T > 0$;
- (iv) $\mathbb{Q}$ is not equivalent to $\mathbb{P}$ on $\mathcal{F}_T$.

Proof. By Definition 4.1 S is a local martingale; hence the strict local martingale property of S is equivalent to $\mathfrak{m}_T = 1 - \mathbb{E}^{\mathbb{Q}}(S_T) > 0$. In view of (4.1) this shows equivalence of the first three assertions. Finally $\frac{d\mathbb{Q}}{d\mathbb{P}}|_{\mathcal{F}_T} = M_T$ and hence $\mathbb{Q}$ is not equivalent to $\mathbb{P}$ if and only if $M_T = 0$ with positive $\mathbb{P}$ -probability, which is exactly assertion (ii). ■

Example 4.4. We now continue Example 2.1, i.e., we consider M given by $dM_t = \sigma(M_t)dW_t^{\mathbb{P}}$, where σ satisfies the same conditions as before. In addition assume $\int_1^\infty x\sigma^{-2}(x)dx = \infty$, such that M is a true $\mathbb{P}$ -martingale by (2.3), and let τ denote the first hitting time of zero of M . Construct now the measure $\mathbb{Q}$ via $\frac{d\mathbb{Q}}{d\mathbb{P}}|_{\mathcal{F}_t} = M_t$, and set $S_t := M_t^{-1}\mathbf{1}_{\{t < \tau\}}$. It is easy to see that S satisfies the stochastic differential equation $dS_t = \tilde{\sigma}(S_t)dW_t^{\mathbb{Q}}$, with $S_0 = 1$, $\tilde{\sigma}(y) = y^2\sigma(1/y)$, and where $W^{\mathbb{Q}}$ is a $\mathbb{Q}$ -Brownian motion defined up to time τ . Note that the following equalities hold:

$$(4.2) \quad \int_0^1 \frac{y}{\tilde{\sigma}^2(y)} dy = \int_1^\infty \frac{x}{\sigma^2(x)} dx \quad \text{and} \quad \int_1^\infty \frac{y}{\tilde{\sigma}^2(y)} dy = \int_0^1 \frac{x}{\sigma^2(x)} dx.$$

The first integral is infinite by assumption, and it follows from the integral conditions (2.2) and (2.3) in Example 2.1 that S is $\mathbb{Q}$ -a.s. strictly positive. If the second integral is infinite it follows from the same conditions that also M is $\mathbb{P}$ -a.s. positive and that S is a true $\mathbb{Q}$ -martingale. If the second integral is finite, then M has mass at zero and S is a strict local martingale. Other cases are not possible, in line with Lemma 4.3. We remark that the duality relations between M and S and between σ and $\tilde{\sigma}$, as well as their relevance for the strict local martingale property of S , have been already described in [16].

Relations between option prices in dual models are well known in the case where both S and M are true martingales without mass at zero [17, 21] and have been described in the strict local martingale case in [34, 44]. We recall the relation for European puts and calls. In addition to the put and call prices P_S and C_S^α described above, we write $P_M(x) := \mathbb{E}^{\mathbb{P}}(e^x - M_T)_+$ and $C_M(x) := \mathbb{E}^{\mathbb{P}}(M_T - e^x)_+$ for the European put and call option on the underlying M under $\mathbb{P}$.

Proposition 4.5. *Let $(M, \mathbb{P})$ and $(S, \mathbb{Q})$ be market models in duality with time horizon $T > 0$. Then for any $x \in \mathbb{R}$ and any $\alpha \in [0, 1]$, the following relations between put and call prices hold:*

$$(4.3) \quad C_S^\alpha(x) = e^x P_M(-x) + (\alpha - 1)\mathfrak{m}_T \quad \text{and} \quad P_S(x) = e^x C_M(-x).$$

Proof. First, note that

$$\mathbb{Q}(T < \tau) = \mathbb{E}^\mathbb{P} [M_T \mathbf{1}_{\{T < \tau\}}] = \mathbb{E}^\mathbb{P} [M_T] = 1.$$

For puts we compute

$$\begin{aligned} P_S(x) &= \mathbb{E}^\mathbb{Q} (e^x - S_T)_+ = \mathbb{E}^\mathbb{Q} [\mathbf{1}_{\{T < \tau\}} (e^x - S_T)_+] = \mathbb{E}^\mathbb{P} \left[M_T \mathbf{1}_{\{T < \tau\}} \left(e^x - \frac{1}{M_T} \right)_+ \right] \\ &= e^x \mathbb{E}^\mathbb{P} [\mathbf{1}_{\{T < \tau\}} (M_T - e^{-x})_+] = e^x \mathbb{E}^\mathbb{P} [(M_T - e^{-x})_+] = e^x C_M(-x), \end{aligned}$$

where we have used that $M_T = 0$, $\mathbb{P}$ -a.s. on $\{\tau \leq T\}$. The results for calls follow from put-call parity for the $\mathbb{P}$ -martingale M and “modified put-call parity” (2.4) for the local $\mathbb{Q}$ -martingale S . ■

4.2. Duality and symmetry of the implied volatility. We consider the consequences of the duality relation studied in section 4.1 on implied volatilities.

Theorem 4.6. *Let S be a strictly positive strict local $\mathbb{Q}$ -martingale in duality with the true $\mathbb{P}$ -martingale M with mass at zero. Denote by $I_M(x)$ the implied volatility under $\mathbb{P}$ for log-strike x and underlying M . Then, for all $x \in \mathbb{R}$,*

$$(4.4) \quad I_S^p(x) = I_S^1(x) = I_M(-x).$$

Proof. For any $x \in \mathbb{R}$, the implied volatility I_S^p is the unique solution to the equation $P_S(x) = P_{BS}(x, I_S^p(x))$. Therefore, using (4.3), we can write

$$\begin{aligned} e^x \mathcal{N}(-d_-(x, I_S^p(x))) - \mathcal{N}(-d_+(x, I_S^p(-x))) &= P_S(x) = e^x C_{BS}(-x, I_M(-x)) \\ &= e^x \mathcal{N}(d_+(-x, I_M(-x))) - \mathcal{N}(d_-(-x, I_M(-x))) \\ &= e^x \mathcal{N}(-d_-(x, I_M(-x))) - \mathcal{N}(-d_+(x, I_M(-x))). \end{aligned}$$

The theorem follows by uniqueness of the implied volatilities $I_S^p(x)$ and $I_M(-x)$. ■

Using the above theorem we can translate between results on the right-wing behavior of the implied volatility in strict local martingale models (such as Theorem 3.3) and results on the left-wing behavior of the implied volatility in true martingale models with mass at zero (as given in [11]). Applying the duality to [11, Theorem 4.2], we obtain the following refined representation of the error term $\varepsilon(x)$ in Theorem 3.3.

Corollary 4.7. *Let S be a càdlàg and strictly positive strict local $\mathbb{Q}$ -martingale, $T > 0$, and $\mathfrak{m}_T$ the martingale defect of S . Set $G(x) := \mathbb{E}^\mathbb{Q}(S_T \mathbf{1}_{\{S_T \geq e^x\}})$ and $\mathfrak{n}_T := \mathcal{N}^{-1}(\mathfrak{m}_T)$. If $G(x) = \mathcal{O}(x^{-3/2})$ as x tends to infinity, then in the put-implied volatility expansion of Theorem 3.3, $\varepsilon(x)$ can be replaced by*

$$\varepsilon(x) = \frac{\frac{1}{2}n_T^2 + 1}{\sqrt{2Tx}} + \frac{n_T}{4x\sqrt{T}} + \mathcal{O}(x^{-3/2}).$$

Proof. Let F denote the cumulative distribution function of M_T under $\mathbb{P}$. Then

$$G(x) = \mathbb{E}^{\mathbb{Q}}(S_T \mathbf{1}_{\{S_T \geq e^x\}}) = \mathbb{E}^{\mathbb{P}}(\mathbf{1}_{\{M_T \leq e^{-x}\}} \mathbf{1}_{\{T < \tau\}}) = F(e^{-x}) - F(0).$$

The remaining claims follow from the duality relation $I_M(-x) = I_S^1(x)$ and [11, Theorem 4.2]. ■

In [11, Corollary A.2], the authors showed that if the underlying is given by a true martingale with mass at zero, then the implied volatility smile cannot be symmetric (in the sense that $I_M(x) = I_M(-x)$). The dual result for strict local martingale models is the following.

Corollary 4.8. *If S is a strictly positive strict local $\mathbb{Q}$ -martingale, then the put smile is not symmetric.*

Remark 4.9. For call-implied volatilities I_S^α , the smile can never be symmetric for $\alpha \in [0, 1)$ for the simple reason that $I_S^\alpha(x)$ does not exist for sufficiently small strikes, by Theorem 3.1. In the fully collateralized case $\alpha = 1$, the call-implied volatility coincides with the put-implied volatility and Corollary 4.8 applies.

To avoid the technical conditions of Definition 4.1 we give a direct proof.

Proof of Corollary 4.8. Since S is strictly positive, we have $\mathbb{Q}(S_T = 0) = 0$. By [11, Theorem 3.6], this implies that $\lim_{x \downarrow -\infty} (\sqrt{T} I_S^p(x) - \sqrt{2|x|}) = -\infty$. Assuming that $I_S^p(x)$ is symmetric, the same expansion has to hold true as x tends to $+\infty$. This is a contradiction to Theorem 3.3 and we conclude that $I_S^p(x)$ cannot be symmetric. ■

Again looking at the example of one-dimensional diffusions is instructive.

Example 4.10. We continue Examples 2.1 and 4.4 and consider the $\mathbb{P}$ -martingale M defined by $dM_t = \sigma(M_t) dW_t^\mathbb{P}$ and the dual $\mathbb{Q}$ -martingale S given by $dS_t = \tilde{\sigma}(S_t) dW_t^\mathbb{Q}$, where $\tilde{\sigma}(y) = y^2 \sigma(1/y)$. The dual models $(S, \mathbb{Q})$ and $(M, \mathbb{P})$ have the same distribution if $\sigma = \tilde{\sigma}$. In this case it follows from (4.2) and the discussion in Example 4.4 that either S and M are both true martingales without mass at zero or they are both strict local martingales with mass at zero. Being a true martingale with mass at zero or a strict local martingale without mass at zero is not compatible with the symmetry assumption $\sigma = \tilde{\sigma}$, in line with [11, Corollary A.2] and Corollary 4.8.

5. Examples.

5.1. Strict local martingale with deterministic endpoint. We borrow this example from [10] and [13]. Let $T > 0$, $\mu \in (0, 1)$, W be a standard Brownian motion, and $(S_t)_{t \geq 0}$ be the unique strong solution to the stochastic differential equation

$$dS_t = \frac{S_t - \mu}{\sqrt{T - t}} dW_t,$$

starting from $S_0 = 1$. On $[0, T)$, it is clear that $S_t = (1 - \mu) \exp(W_{\varphi_t} - \frac{1}{2}\varphi_t) + \mu$, where $\varphi_t \equiv -\log(1 - t/T)$, so that for any $\mu \in \mathbb{R}$, S is a true martingale on $[0, T)$, but $S_T = \mu$ almost surely. Furthermore, since $\mathbf{m}_T = 1 - \mathbb{E}^\mathbb{Q}(S_T) = 1 - \mu$, the call and put prices read $C_S(x) = (\mu - e^x)_+ = (1 - \mathbf{m}_T - e^x)_+$ and $P_S(x) = (e^x - \mu)_+ = (e^x - 1 + \mathbf{m}_T)_+$, so that the lower bounds in (2.5) and (2.6) are attained. Setting $\mu = 0$ we can create a strict local martingale

with mass at zero at time T , i.e., an example of a process that does not fall within the scope of the duality approach.

5.2. The constant elasticity of variance model. On a probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{Q})$ supporting a Brownian motion W , the stochastic differential equation $dS_t = \sigma S_t^{1+\beta} dW_t$, with $S_0 > 0$, admits a unique strong solution, which is a true martingale if and only if $\beta \leq 0$ (see [31, Chapter 6.4]). In the case $\beta > 0$, S is a strict local martingale and the process $Y := S^{-\beta}/(\sigma\beta)$ is a Bessel process of index $1/(2\beta)$ (equivalently of dimension $2 + 1/\beta$). By (2.2), the origin is unattainable in finite time, and for any $t > 0$, the transition density of S_t reads

$$\mathbb{P}(S_t \in ds) = \frac{S_0^{1/2} s^{-2\beta-3/2}}{\sigma^2 \beta t} \exp\left(-\frac{S_0^{-2\beta} + s^{-2\beta}}{2\sigma^2 \beta^2 t}\right) I_{1/(2\beta)}\left(\frac{(S_0 s)^{-\beta}}{\sigma^2 \beta^2 t}\right) ds,$$

where $I(\cdot)$ is the modified Bessel function of the first kind. The expectation can be computed in closed form as $\mathbb{E}^{\mathbb{Q}}(S_t) = S_0 \Gamma(-\frac{1}{2\beta}, \frac{S_0^{-2\beta}}{2\sigma^2 \beta^2 t})$, where $\Gamma(a, x) := \Gamma(a)^{-1} \int_0^x u^{a-1} e^{-u} du$ is the (normalized) incomplete Gamma function. In the case $\beta = 1$, the density simplifies to

$$(5.1) \quad \mathbb{P}(S_t \in ds) = \frac{S_0}{\sigma s^3 \sqrt{2\pi t}} \left\{ \exp\left(-\frac{(1/s - 1/S_0)^2}{2t\sigma^2}\right) - \exp\left(-\frac{(1/s + 1/S_0)^2}{2t\sigma^2}\right) \right\} ds,$$

and $\mathbb{E}(S_T) = S_0(1 - 2\mathcal{N}(-1/(S_0\sqrt{T})))$. In Figure 1, we illustrate numerically Theorem 3.3, including the higher-order terms of the refined expansion in Corollary 4.7. First, we need to verify the assumptions on the behavior of the complementary cumulative distribution function $G(x) = \int_{e^x}^{\infty} s\mathbb{P}(S_t \in ds)$ as x tends to infinity. As s becomes large, the modified Bessel function behaves, for ν not a negative integer, as [1, section 9.6]

$$I_{\nu}(z) = \left(\frac{z}{2}\right)^{\nu} \sum_{n \geq 0} \frac{1}{n! \Gamma(\nu + n + 1)} \left(\frac{z^2}{4}\right)^n.$$

Therefore, since $\beta > 0$ and $1/(2\beta)$ is clearly not a negative integer, as s tends to infinity,

$$\begin{aligned} \mathbb{P}(S_t \in ds) &\sim \frac{S_0^{1/2} s^{-2\beta-3/2}}{\sigma^2 \beta t} \exp\left(-\frac{S_0^{-2\beta} + s^{-2\beta}}{2\sigma^2 \beta^2 t}\right) \left(\frac{(S_0 s)^{-\beta}}{2\sigma^2 \beta^2 t}\right)^{1/(2\beta)} ds, \\ &\sim c_0 s^{-2(\beta+1)} ds, \end{aligned}$$

for some $c_0 > 0$, depending on S_0, t, σ, β . Karamata's theorem [6, Theorem 2.6.5] implies, since $\beta > 0$, that

$$\int_{e^x}^{\infty} s\mathbb{P}(S_t \in ds) \sim \frac{c_0}{2\beta} e^{-2\beta x},$$

as x tends to infinity, ensuring that the assumptions in Corollary 4.7 hold. Regarding the large-time behavior of the put-implied volatility smile, Theorem 3.8 yields (see also [46, Example 5.9])

$$I_S(x, T)^2 T = 4 \log(T) - 4 \log \log(T) + 4x + \varepsilon(x, T)$$

for all $x \in \mathbb{R}$ as T tends to infinity.

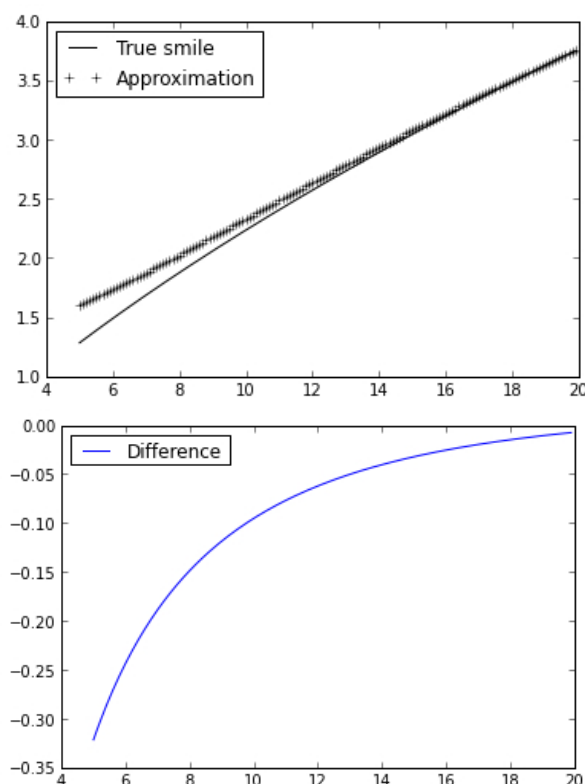


Figure 1. Implied volatility and its approximation in the constant elasticity of variance model with parameters $(S_0, \beta, \sigma, T) = (1, 2.4, 10\%, 1)$. The horizontal axis shows log-strike; in the first subplot the true value of $x \mapsto I_S^p(x)$ (solid line) and its approximation from Theorem 3.3 including the higher-order terms of Corollary 4.7 (crosses) are shown. The second subplot shows the error between the true value and its approximation.

Acknowledgment. Both authors would like to thank two anonymous referees for their valuable comments, which have led to significant improvements of our manuscript.

REFERENCES

- [1] M. ABRAMOWITZ AND I. STEGUN, *Handbook of Mathematical Functions: With Formulas, Graphs, and Mathematical Tables*, National Bureau of Standards, 1964.
- [2] L. ANDERSEN, *Option pricing with quadratic volatility: A revisit*, *Finance Stoch.*, 15 (2011), pp. 191–219.
- [3] J. AZÉMA, R. F. GUNDY, AND M. YOR, *Sur l'intégrabilité uniforme des martingales continues*, *Sémin. Probab.*, 14 (1980), pp. 53–61.
- [4] S. BENAÏM AND P. K. FRIZ, *Regular variation and smile asymptotics*, *Math. Finance*, 19 (2009), pp. 1–12.
- [5] S. BENAÏM, P. K. FRIZ, AND R. LEE, *On the Black–Scholes Implied Volatility at Extreme Strikes*, in *Frontiers in Quantitative Finance: Volatility and Credit Risk Modeling*, Wiley, New York, 2008.
- [6] N. BINGHAM, C. GOLDIE, AND J. TEUGELS, *Regular Variation*, Cambridge University Press, Cambridge, UK, 1987.
- [7] T. BJÖRK, *Arbitrage Theory in Continuous Time*, Oxford University Press, New York, 2004.
- [8] S. BLEI AND H. J. ENGELBERT, *On exponential local martingales associated with strong Markov continuous local martingales*, *Stochastic Process. Appl.*, 119 (2009), pp. 2859–2880.

- [9] P. CARR, J. RUF, AND T. FISHER, *On the hedging of options on exploding exchange rates*, Finance Stoch., 18 (2014), pp. 115–144.
- [10] A. COX AND D. HOBSON, *Local martingales, bubbles and option prices*, Finance Stoch., 9 (2005), pp. 477–492.
- [11] S. DE MARCO, C. HILLAIRET, AND A. JACQUIER, *Shapes of implied volatility with positive mass at zero*, SIAM J. Financial Math., 8 (2017), pp. 709–737.
- [12] F. DELBAEN AND W. SCHACHERMAYER, *A general version of the fundamental theorem of asset pricing*, Math. Ann., 300 (1994), pp. 463–520.
- [13] F. DELBAEN AND W. SCHACHERMAYER, *Arbitrage possibilities in Bessel processes and their relations to local martingales*, Probab. Theory Related Fields, 102 (1995), pp. 357–366.
- [14] F. DELBAEN AND W. SCHACHERMAYER, *The fundamental theorem of asset pricing for unbounded stochastic processes*, Math. Ann., 312 (1998), pp. 215–250.
- [15] F. DELBAEN AND W. SCHACHERMAYER, *The Mathematics of Arbitrage*, Springer Finance, New York, 2006.
- [16] F. DELBAEN AND H. SHIRAKAWA, *No arbitrage condition for positive diffusion price processes*, Asia-Pac. Financ. Markets, 9 (2002), pp. 159–168.
- [17] E. EBERLEIN, A. PAPAPANTOLEON, AND A. N. SHIRYAEV, *On the duality principle in option pricing: Semimartingale setting*, Finance Stoch., 2 (2008), pp. 265–292.
- [18] H.-J. ENGELBERT AND T. SENF, *On functionals of Wiener process with drift and exponential local martingales*, in Stochastic Processes and Related Topics, Trends in Math., Springer, New York, 1990, pp. 45–58.
- [19] H. FÖLLMER, *The exit measure of a supermartingale*, Z. Wahrscheinlichkeitstheorie Verwandte Gebiete, 21 (1972), pp. 154–166.
- [20] J. GATHERAL, *The Volatility Surface*, Wiley, New York, 2006.
- [21] H. GEMAN, N. EL KAROUI, AND J.-C. ROCHET, *Changes of numeraire, changes of probability measure and option pricing*, J. Appl. Probab., 32 (1995), pp. 443–458.
- [22] A. GULISASHVILI, *Left-wing asymptotics of the implied volatility in the presence of atoms*, Int. J. Theor. Appl. Finance, 18 (2015).
- [23] A. GUSHCHIN, M. URUSOV, AND M. ZERVOS, *On the submartingale/supermartingale property of diffusions in natural scale*, Proc. Steklov Inst. Math., 287 (2014), pp. 122–132.
- [24] S. HESTON, M. LOEWENSTEIN, AND G. WILLARD, *Options and bubbles*, Rev. Financ. Stud., 20 (2007), pp. 359–390.
- [25] D. HOBSON, *Comparison results for stochastic volatility models via coupling*, Finance Stoch., 14 (2010), pp. 129–152.
- [26] H. HULLEY AND E. PLATEN, *A visual classification of local martingales*, in Seminar on Stochastic Analysis, Random Fields and Applications, Progr. Probab., Birkhäuser/Springer, Basel, 2011, pp. 147–157.
- [27] J. JACOD AND A. N. SHIRYAEV, *Limit Theorems for Stochastic Processes*, 2nd ed., Springer, Berlin, 2003.
- [28] R. JARROW, Y. KCHIA, AND P. PROTTER, *How to detect an asset bubble*, SIAM J. Financial Math., 2 (2011), pp. 839–865.
- [29] R. JARROW, Y. KCHIA, AND P. PROTTER, *Is there a bubble in LinkedIn’s stock price?*, J. Portfolio Management, 38 (2011), pp. 125–130.
- [30] R. JARROW, Y. KCHIA, AND P. PROTTER, *Is gold in a bubble?*, Bloomberg’s Risk Newsletter, 26 (2011), pp. 8–9.
- [31] M. JEANBLANC, M. YOR, AND M. CHESNEY, *Mathematical Methods for Financial Markets*, Springer, New York, 2009.
- [32] B. JOURDAIN, *Loss of Martingality in Asset Price Models with Lognormal Stochastic Volatility*, preprint, CERMICS, Paris, 2004.
- [33] I. KARATZAS AND S. E. SHREVE, *Brownian Motion and Stochastic Calculus*, 2nd ed., Springer, New York, 1998.
- [34] C. KARDARAS, D. KREHER, AND A. NIKEGBALI, *Strict local martingales and bubbles*, Ann. Appl. Probab., 25 (2015), pp. 1827–1867.
- [35] R. LEE, *The moment formula for implied volatility at extreme strikes*, Math. Finance, 14 (2004), pp. 469–480.
- [36] A. LEWIS, *Option Valuation Under Stochastic Volatility*, Finance Press, Newport Beach, CA, 2000.

- [37] M. LOEWENSTEIN AND G. WILLARD, *Rational equilibrium asset-pricing bubbles in continuous trading models*, J. Econom. Theory, 91 (2000), pp. 17–58.
- [38] D. MADAN AND M. YOR, *Itô's integrated formula for strict local martingales*, in Séminaire de Probabilités, 39, Lecture Notes in Math. 1874, Springer, New York, 2006, pp. 157–170.
- [39] P. A. MEYER, *La mesure de H. Föllmer en théorie des surmartingales*, Sémin. Probab. Strasbourg, 6 (1972), pp. 118–129.
- [40] S. PAL AND P. PROTTER, *Analysis of continuous strict local martingales via h-transforms*, Stochastic Process. Appl., 120 (2010), pp. 1424–1443.
- [41] P. PROTTER, *A mathematical theory of financial bubbles*, in Paris-Princeton Lectures on Mathematical Finance, Lecture Notes in Math. 2081, Springer, New York, 2013, pp. 1–108.
- [42] L. C. G. ROGERS AND M. TEHRANCHI, *Can the implied volatility move by parallel shifts*, Finance Stochastics, 14 (2010), pp. 235–248.
- [43] L. C. G. ROGERS AND D. WILLIAMS, *Diffusions, Markov Processes, and Martingales: Volume 1, Foundations*, 2nd ed., Cambridge University Press, Cambridge, UK, 2000.
- [44] J. RUF, *Hedging under arbitrage*, Math. Finance, 23 (2013), pp. 297–317.
- [45] C. SIN, *Complications with stochastic volatility models*, Adv. Appl. Probab., 30 (1998), pp. 256–268.
- [46] M. TEHRANCHI, *Asymptotics of implied volatility far from maturity*, J. Appl. Probab., 46 (2009), pp. 629–650.