PATH-DEPENDENT VOLATILITY

JULIEN GUYON
QUANTITATIVE RESEARCH, BLOOMBERG L.P.

Abstract. So far, path-dependent volatility models have drawn little attention from both practitioners and academics compared to local volatility and stochastic volatility models. This is unfair: in this article we show that they combine benefits from both. Like the local volatility model, they are complete and can fit exactly the market smile; like stochastic volatility models, they can produce rich implied volatility dynamics. Not only that: given their huge flexibility, they can actually generate a much broader range of spot-vol dynamics, thus possibly preventing large mispricings, and they can also capture prominent historical patterns of volatility. We give many examples to showcase their capabilities.

1. Introduction

In the field of volatility modeling, three main families of models have been used so far in the finance industry: constant volatility (Bachelier, 1905; Black and Scholes, 1973); local volatility (LV) (Dupire, 1994 [4]); and stochastic volatility (SV) (Hull and White, 1987 [17]; Heston, 1993 [13]; Bergomi, 2005 [1]; among many others). The first two families of models are complete: since the asset price is driven by a single Brownian motion, say W, every payoff admits a unique self-financing replicating portfolio consisting of cash and the underlying asset, so its price is uniquely defined as the initial value of the replicating portfolio, independently of utilities or preferences. Unlike the constant volatility models, the LV model is flexible enough to fit any arbitrage-free surface of implied volatilities (henceforth, 'smile')—but then no more flexibility is left. Calibrating to the market smile is useful when one sells an exotic option whose risk is well mitigated by trading vanilla options—then the model correctly prices the hedging instruments at inception.

For their part, SV models are incomplete: the volatility is driven by one or several extra Brownian motions, not perfectly correlated with W, and as a result perfect replication and price uniqueness are lost. Modifying the drift of the SV leaves the model arbitrage-free, but changes option prices. Several concepts have been suggested to define the price of an option in incomplete models, including super-replication, indifference pricing, quantile hedging, and minimum variance hedging; for a review of these, see e.g. [10].

The main reason to consider SV models is to gain control on key risk factors like volatility of volatility ('vol of vol'), forward skew, and spot-vol correlation. For instance, using a very large mean reversion together with a large vol of vol and a very negative spot-vol correlation, one can generate an almost flat implied volatility surface, together with very negative short term forward skews. If an LV model were used to match this smile, the LV surface would be almost flat as well, producing vanishing forward skew. As a result, cliquets of forward starting call spreads would be much cheaper in the LV model. This is still true even if the smile is not flat: LV models typically underprice these options—we say, underprice the forward skew. As for the vol of vol, controling it avoids mispricing options on realized variance, or forward starting calls, for instance. To sum up, SV models generate joint dynamics of the asset and its implied volatilities (spot-vol dynamics henceforth) that are much richer than the LV ones.

To allow SV models to perfectly calibrate to the market smile, one can use mixed stochastic local volatility (SLV) models, i.e., multiply the SV by an LV (the so-called 'leverage function') which is fitted to the smile using the particle method; see [9]. This modifies the spot-vol dynamics, but rather slightly: usually the leverage function, seen as a function of the asset price, becomes flatter and flatter as time t grows, so the SLV dynamics become closer and closer to pure SV ones as time passes, and the model still generates large vol of vol and large forward skew, except maybe close to $t = 0$.

Date: First version: February 28, 2014. This version: September 10, 2014.
Key words and phrases. Option pricing, path-dependent volatility, complete models, smile calibration, particle method, implied volatility dynamics, ARCH/GARCH models.

Electronic copy available at: http://ssrn.com/abstract=2425048

At this point a natural question arises: can we build *complete* models that have all the nice properties of SLV models, namely, rich spot-vol dynamics, and calibration to the market smile? For instance, can we build a complete model that is calibrated to a flat smile, and yet produces very negative short term forward skews? It is tempting but wrong to quickly answer 'no', by arguing that the only complete model calibrated to the smile is the LV model. This is *not* true: in this article, we will show that path-dependent volatility (PDV) models, which are complete, can produce rich spot-vol dynamics and, furthermore, can be perfectly calibrated to the market smile. The two main benefits of model completeness are price uniqueness and parsimony: due to the causal link in the model between today's asset price movements and future volatility, delta-hedging perfectly replicates the latter—and any payoff; and even if perfect delta-hedging is deemed unrealistic, it is remarkable that so many popular properties of SLV models can be captured using a single Brownian motion. Not only that: we will see that thanks to their huge flexibility, PDV models can generate spot-vol dynamics that are not attainable by SLV models. Section 2 introduces the class of PDV models. Then in Section 3 we explain how we calibrate them to the market smile, before we investigate and illustrate with many numerical examples how to pick a particular PDV in Sections 4 and 5.

2. Path-dependent volatility models

PDV models are those models where the instantaneous volatility σ_t depends on the *path* followed by the asset price so far:

$$\frac{dS_t}{S_t} = \sigma(t, (S_u, u \le t))\, dW_t$$

where, for simplicity, we have taken zero interest rates, repo, and dividends. In practice, the volatility $\sigma_t \equiv \sigma(t, S_t, X_t)$ will often be assumed to depend on the path only through the current value S_t and a finite set X_t of path-dependent variables, which may include for example running or moving averages, maximums or minimums, realized variances, etc.

PDV models have been widely overlooked, compared to LV and SV models. The most famous PDV models are probably the ARCH model by Engle [5] and its descendants GARCH [2], NGARCH, IGARCH, etc. But these are discrete-time models which are hardly used in the derivatives industry. The two other main contributions so far are due to Hobson and Rogers [15] and Bergomi [1]. The discrete setting version of Bergomi's SV model is actually a mixed SV-PDV model in which, given a realization of the variance swap volatility at time $T_i = i\Delta$ for maturity T_{i+1}, $\sqrt{\xi^i_{T_i}}$, the (continuous time) volatility of the underlying on $[T_i, T_{i+1}]$ is path-dependent: it reads $\sigma(S_t/S_{T_i})$, where σ is calibrated to both $\xi^i_{T_i}$ and a desired value of the forward at-the-money (ATM) skew for maturity Δ. By restriking S at T_i, the distribution of $S_{T_{i+1}}/S_{T_i}$ is made independent of S_{T_i}, which allows to decouple the short term forward skew and the spot/volatility correlation.

By contrast, the Hobson-Rogers model is a pure PDV model in which the volatility $\sigma_t = \sigma(X_t)$ is a deterministic function of $X_t = (X^1_t, \dots, X^n_t)$, where the X^m_t are exponentially weighted moments of all the past log increments of the asset price:

$$X^m_t = \int_{-\infty}^t \lambda e^{-\lambda(t-u)} \left(\ln \frac{S_t}{S_u}\right)^m du$$

Consider the case where $n = 1$: then the authors stipulate that the volatility depends only on the offset $X^1_t = \ln S_t - \int_{-\infty}^t \lambda e^{-\lambda(t-u)} \ln S_u\, du$, which is the difference between the current log price and a weighted average of past log prices: the volatility is completely determined by the local trend of the asset price over a period of order $1/\lambda$ years (e.g., 1 month if $\lambda = 12$). As we will see in Section 5, this is supported by empirical studies. Here the choice of an infinite time window and exponential weights is only guided by computational convenience: it ensures that (S_t, X_t) is a Markovian process, so the time-t price of a European payoff of the type $g(S_T, X_T)$—in particular the price of a vanilla option—reads $u(t, S_t, X_t)$ where u is the solution to a second order parabolic partial differential equation. Note in particular that the implied volatilities at time 0 in the model depend not only on the strike, maturity, and S_0, but also on all the past asset prices through X_0.

At this point four natural questions arise:

(1) Can we specify $\sigma(\cdot)$ and λ so as to exactly fit the market smile? [19, 6] only gave approximate calibration results.

Electronic copy available at: http://ssrn.com/abstract=2425048

(2) Does the calibrated model have desired dynamics of implied volatility, such as large negative short term forward skew for instance?
(3) In the definition of X_t, can we use general weights and a finite time window $[t-\Delta, t]$ instead of $(-\infty, t]$, so that the volatility truly depends only a limited portion of the past, e.g., the previous month? The generalization in [7] is very partial as it requires positive weights on $[0, t]$.
(4) Much more importantly: how do we generalize to other choices of X_t? The generalization in [16], where the volatility depends on a particular modified version of the offset X_t^1, is also very partial.

In this article we solve all these questions at a time: first we choose *any* set of path-dependent variables X_t and *any* function $\sigma(t, S, X)$ so that the PDV model with $\sigma_t = \sigma(t, S_t, X_t)$ has desired spot-vol dynamics and/or captures historical features of volatility, and then we define a new PDV model by multiplying $\sigma(t, S, X)$ by a leverage function $l(t, S)$ and we perfectly calibrate l to the market smile of S using the particle method. When we do not calibrate to the smile ($l \equiv 1$), we speak of 'pure' PDV.[1] Usually, multiplying the pure PDV $\sigma(t, S_t, X_t)$ by the calibrated leverage function distorts only slightly the spot-vol dynamics since leverage functions typically flatten over time (see Figure 4.3). This way we mimic SLV models, with the pure PDV $\sigma(t, S_t, X_t)$ playing the role of the SV, but we stay in the world of complete models.

Actually, the same program can be run by choosing two functions $a(t, S, X)$ and $b(t, S, X)$ instead of only one function $\sigma(t, S, X)$, and then defining $\sigma_t^2 = a(t, S_t, X_t) + b(t, S_t, X_t) l(t, S_t)$. The case where $b \equiv 1$ is the complete analogue of incomplete *additive* SLV models, where the instantaneous variance is the *sum* of an SV and an LV.

3. Smile calibration of path-dependent volatility models

Before investigating how to choose X and $\sigma(t, S, X)$, let us explain how, given such a choice, we can uniquely build the leverage function $l(t, S)$ such that the PDV model

$$\frac{dS_t}{S_t} = \sigma(t, S_t, X_t) l(t, S_t)\, dW_t \tag{3.1}$$

fits exactly the market smile of S. From Itô-Tanaka's formula—or, in this deterministic interest rate framework, from Gyöngy's theorem—we know that Model (3.1) is exactly calibrated to the market smile of S if and only if

$$\mathbb{E}^{\mathbb{Q}}[\sigma(t, S_t, X_t)^2 | S_t] l(t, S_t)^2 = \sigma_{\mathrm{Dup}}^2(t, S_t) \tag{3.2}$$

where $\mathbb{Q}$ denotes the unique risk-neutral measure and σ_{Dup} the Dupire LV. As a consequence, the calibrated model satisfies the nonlinear McKean stochastic differential equation

$$\frac{dS_t}{S_t} = \frac{\sigma(t, S_t, X_t)}{\sqrt{\mathbb{E}^{\mathbb{Q}}[\sigma(t, S_t, X_t)^2 | S_t]}} \sigma_{\mathrm{Dup}}(t, S_t)\, dW_t \tag{3.3}$$

The particle method (see [9]) is a very efficient and elegant Monte Carlo method that computes the above conditional expectation, hence the leverage function $l(t, S) = \sigma_{\mathrm{Dup}}(t, S) / \sqrt{\mathbb{E}^{\mathbb{Q}}[\sigma(t, S_t, X_t)^2 | S_t = S]}$, on the go while simulating the paths.

Note that for many choices of X_t it should be possible in theory to choose a PDV $\sigma(t, X_t)$ with no explicit dependence on S_t such that the smile calibration condition $\mathbb{E}^{\mathbb{Q}}[\sigma(t, X_t)^2 | S_t] = \sigma_{\mathrm{Dup}}^2(t, S_t)$ is satisfied—there are enough degrees of freedom in σ to handle the double infinity of constraints indexed by (t, S). However so far we do not know how to *build* such σ. Multiplying a given $\sigma(t, X)$ by a leverage function $l(t, S)$ allows us to build a calibrating PDV $\sigma(t, X) l(t, S)$, but to the expense of introducing an explicit dependence on S.

In a recent paper [3] Brunick and Shreve have shown how, under very mild assumptions, given a general Itô process $dS_t = \sigma_t S_t\, dW_t$ and a special type of path-dependent variable X, one can build a PDV $\sigma(t, S_t, X_t)$ such that, for each t, the *joint* distribution of (S_t, X_t) is the same in both models. A solution is shown to be $\sigma(t, S_t, X_t)^2 = \mathbb{E}^{\mathbb{Q}}[\sigma_t^2 | S_t, X_t]$. However, only X's satisfying a type of Markov property are admissible. For instance, running averages and running minimums/maximums are admissible, but a moving average is not; instead, one must pick $X_t = (S_u, t - \Delta \le u \le t)$, where Δ is the length of the moving time window. If X is admissible, and if a formula is available to compute $\sigma(t, S_t, X_t)$ from market prices of options, the

[1] Alternatively, to mimick the SV/SLV classification, we could speak of path-dependent local volatility (PDLV) for the calibrated model (3.1), and reserve the use of 'PDV' for the cases where $l \equiv 1$. However, this is ambiguous, as a PDLV can be seen as a PDV, since $\sigma(t, S, X) l(t, S) \equiv \Sigma(t, S, X)$ is a function of the asset path.

X_t	$\sigma(S,X)$ producing large forward skew		$\sigma(S,X)$ producing U-shaped forward smile	
$S_{t-\Delta}$	1	$\overline{\sigma}1_{\left\{\frac{S}{X}\le 1\right\}}+\underline{\sigma}1_{\left\{\frac{S}{X}>1\right\}}$	4	$\overline{\sigma}1_{\left\{\left\lvert\frac{S}{X}-1\right\rvert>\kappa\sigma_0\sqrt{\Delta}\right\}}+\underline{\sigma}1_{\left\{\left\lvert\frac{S}{X}-1\right\rvert\le\kappa\sigma_0\sqrt{\Delta}\right\}}$
$\overline{S}_t^\Delta$	2	as above	5	as above
(m_t^Δ, M_t^Δ)	3	$\overline{\sigma}1_{\left\{\frac{S-m}{M-m}\le\frac{1}{2}\right\}}+\underline{\sigma}1_{\left\{\frac{S-m}{M-m}>\frac{1}{2}\right\}}$	6	$\overline{\sigma}1_{\left\{\frac{M}{m}-1>\kappa\sigma_0\sqrt{\Delta}\right\}}+\underline{\sigma}1_{\left\{\frac{M}{m}-1\le\kappa\sigma_0\sqrt{\Delta}\right\}}$

TABLE 1. Examples of PDV functions that generate a large negative short term forward skew even when calibrated to a flat surface of implied volatilities.

Brunick-Shreve result, combined with the particle method, allows us to calibrate a diffusive model to those market prices. Henry-Labordère [12] has given an example of such formula where X_t is the running maximum M_t, using barrier option payoffs $(S_T-K)_+1_{M_T>B}$.

Note that, by taking X_t to be the whole path of S itself up to time t, the Brunick-Shreve result shows that the price process produced by any SV/SLV model has the same distribution, *as a process*, as a PDV model. This means that there always exists a PDV model that produces exactly the same prices of, not only vanilla options, but *all* options, including path-dependent, exotic options. It then comes as no surprise that PDV models can reproduce popular SLV spot-vol dynamics (see Section 4). All of them can actually be reproduced by PDV models.

Now the crucial question is: How to choose a particular PDV? There are two main possible goals: (1) to generate desired spot-vol dynamics, and/or (2) to capture historical features of volatility. These two goals are not antagonist: it might very well happen, and it is desirable, that a given choice of a PDV fulfills both objectives at a time. The next two sections investigate these two objectives.

4. Choose a particular PDV to generate desired spot-vol dynamics

Can we choose a PDV $\sigma(t,S,X)$ that generates, for instance, large negative short term forward skews, even when it is calibrated to a flat smile? Of course the flat smile case is unrealistic, but it helps understand the decoupling of the spot smile and the forward smile. Using an analogy with SLV models, we need $\sigma(t,S_t,X_t)$ to be negatively correlated with S_t. Obviously this may be achieved by picking a decreasing function σ of S alone, but in view of (3.3), after 'decorating' the model with a leverage function and calibrating it to the smile, this would bring us back to the pure LV model. Therefore what we actually need is $\sqrt{\eta(t,S_t,X_t)}$ to be negatively correlated with S_t, where

$$\eta(t,S,X)\equiv\frac{\sigma(t,S,X)^2}{\mathbb{E}^{\mathbb{Q}}[\sigma(t,S_t,X_t)^2|S_t=S]}$$

denotes what we call the 'path-dependent to local variance ratio' (PDLVR). The PDLVR, or alternatively its conditional variance

$$D(t,S)=\mathrm{Var}(\eta(t,S_t,X_t)|S_t=S)=\mathbb{E}[(\eta(t,S_t,X_t)-1)^2|S_t=S]$$

measure the deviation from LV, which is characterized by $\eta\equiv 1$, or $D\equiv 0$. Note that the PDLVR itself is uncorrelated with the asset price, since $\mathbb{E}^{\mathbb{Q}}[\eta(t,S_t,X_t)|S_t=S]\equiv 1$. For its square root to be negatively correlated with S_t, $\sigma(t,S_t,X_t)$ must be negatively correlated to S_t, but not perfectly: good PDV candidates must satisfy that the correlation between S_t and $\sigma(t,S_t,X_t)$ is more around, say, -50% than around -1% or -99%. We call 'moderate correlation' this property that, for the calibrated PDV model to produce large negative short term forward skews, the *levels* of the asset price and the PDV must be negatively correlated, but not too much. Note that for all usual SLV models, the levels of the asset price and the SV are moderately negatively correlated, even if their *increments* are extremely negatively correlated, due to the presence of mean reversion in the SV.

Three adequate simple choices of PDV are listed in Table 1 (Examples 1–3), where

$$\overline{S}_t^\Delta=\frac{\int_0^\Delta w_\tau S_{t-\tau}\,d\tau}{\int_0^\Delta w_\tau\,d\tau},\qquad m_t^\Delta=\inf_{t-\Delta\le u\le t}S_u,\qquad\text{and}\qquad M_t^\Delta=\sup_{t-\Delta\le u\le t}S_u$$

respectively denote a weighted moving average, minimum, and maximum of the asset price over the previous Δ years. (Note that, in our numerical experiments, $w_\tau\equiv 1$, and, as a simplifying assumption, we use $\min(\Delta,t)$ instead of Δ so the option prices do not depend on the asset prices for $t<0$.) Let us first look

at Example 1: The PDV takes two values $\underline{\sigma} < \overline{\sigma}$, depending on whether S_t is larger or smaller than $S_{t-\Delta}$. Comparing with SV models, the spread $\overline{\sigma} - \underline{\sigma}$ mainly plays the part of the vol of vol: we need it to be large enough to generate large negative short term forward skew. The length Δ of the time window plays the role of spot-vol correlation because when Δ is very small, the volatility $\sigma(S_t, X_t)$ is almost independent of S_t, as in this case the fact that S_t is small or large gives almost no information on $S_t/S_{t-\Delta}$, whereas when Δ is large enough, if S_t is small, it makes it more likely that S_t be smaller than $S_{t-\Delta}$, which creates a negative spot-vol correlation. However, if Δ is too large, for instance if the time window is the full interval $[0, t]$, then the volatility depends only on S and we are back to the LV model. The 'correlation' parameter Δ must not be too small nor too large, which illustrates the moderate correlation property. But somehow Δ also plays the part of mean reversion in the sense that the smaller Δ, the more ergodic the volatility, hence the flatter the forward smile; therefore a small Δ and a large mean-reversion have analogous effects (see [8]). Example 2 is very similar to Example 1, with $S_{t-\Delta}$ replaced by a moving average $\overline{S}_t^{\Delta}$, which makes more financial sense—why put all the weight w_τ on $\tau = \Delta$? As for Example 3, it uses the fact that $\frac{S_t - m_t^{\Delta}}{M_t^{\Delta} - m_t^{\Delta}}$ is positively correlated with S_t, and the larger Δ, the larger the correlation.

Figure 4.1 compares the prices of forward starting ATM call spreads and digital calls in different models: LV, the PDV models of Examples 1–3 (with their leverage function l calibrated to a flat 20% smile), and the corresponding pure PDV models. It proves that PDV models can be calibrated to a flat smile and generate highly skewed forward implied volatilities. Note that for month #1, the call spread (resp. the digital call) is actually not forward starting: it is a vanilla option, and since the LV and the three PDV models are all calibrated to the same smile, the four prices coincide. As the forward starting date increases, the PDV price gets closer to the pure PDV price, a sign that the leverage function l is flattening around 1. This is confirmed by Figure 4.3 (top left), where we have graphed the leverage function surface $l(t, S)$ of Example 3 (the graphs for Examples 1 and 2 are very similar). In this pedagogic example of a flat smile, l somehow fights against the pure PDV so as to fit the flat smile: it increases with the asset price, sharply for small t. It modifies the model dynamics in the short term. Yet Figure 4.1 shows that, even in this quite extreme toy case, the impact of l on the forward skew vanishes quickly. In realistic cases, it is likely that l be almost flat, or generate a skew in the same direction as the pure PDV model, like for SLV models (see Figure 5.4). Note that the values of $\underline{\sigma}$ and $\overline{\sigma}$ were chosen so that the smiles of the pure PDV models take values around the flat inputed smile (see Figure 4.2, right).

We can use the calibrated SLV model where the SV is modeled as the exponential of an Ornstein-Uhlenbeck process to reproduce the prices of call spreads and digital calls of Example 2 in Figure 4.1. One must then use a very large vol of vol, beyond 550%, together with an extreme spot-vol correlation (say, −99%). This is because we need a *huge* positive *return* in volatility when asset returns are negative. But (a) this creates numerical problems, and (b) a very large mean reversion above 20 is then somehow artificially needed to keep the volatility within a reasonable range. By comparison PDV models, which can directly relate the asset returns to the volatility *levels*, look much more handy and can more naturally generate large negative forward skews.

Obviously one can generate positive skews by exchanging $\overline{\sigma}$ and $\underline{\sigma}$ in the definition of the PDV. Also, one may use smoothed versions of the PDV by replacing the Heaviside function by $\frac{1}{2}(1 + \tanh(\lambda x))$ for instance. However, the fact that in Examples 1–3 the instantaneous volatility is a jump process is not problematic: no one has ever seen such quantity—it may actually not exist. Figure 5.1 shows that a two-state instantaneous volatility can accurately capture historically observed patterns of short term ATM implied volatilities.

Now, what if we want a PDV model calibrated to a flat smile and yet that generates pronounced U-shaped short term forward smiles? Inspired again by known results on SLV models, we need $\sqrt{\eta(t, S_t, X_t)}$ to be highly volatile and uncorrelated with S_t. However, this is not sufficient: we also need that $\sqrt{\eta(t, S_t, X_t)}$ be not ergodic—at least at the scale of the maturity τ of the forward smile considered, e.g., $\tau = 1/12$ (1 month). This means that if we used Examples 1–3 of Table 1, we would need to choose $\Delta \approx \tau$ or $\Delta \gg \tau$. Now, $\Delta \approx \tau$ implies that $\sqrt{\eta(t, S_t, X_t)}$ is correlated with S_t; and if Δ is too large, then $\sqrt{\eta(t, S_t, X_t)}$ is almost constant, hence not volatile. In both cases this will fail to produce a U-shaped forward smile, so we must turn to new PDVs. Examples 4–6 in Table 1 are natural candidates, for which the volatility is large if and only if so are recent asset returns (up or down). The corresponding prices of forward starting butterfly spreads are shown in Figure 4.2 (left). These models produce vanishing ATM forward skew. The leverage function of Example 6 is reported in Figure 4.3 (top right); those of Examples 4 and 5 are similar.

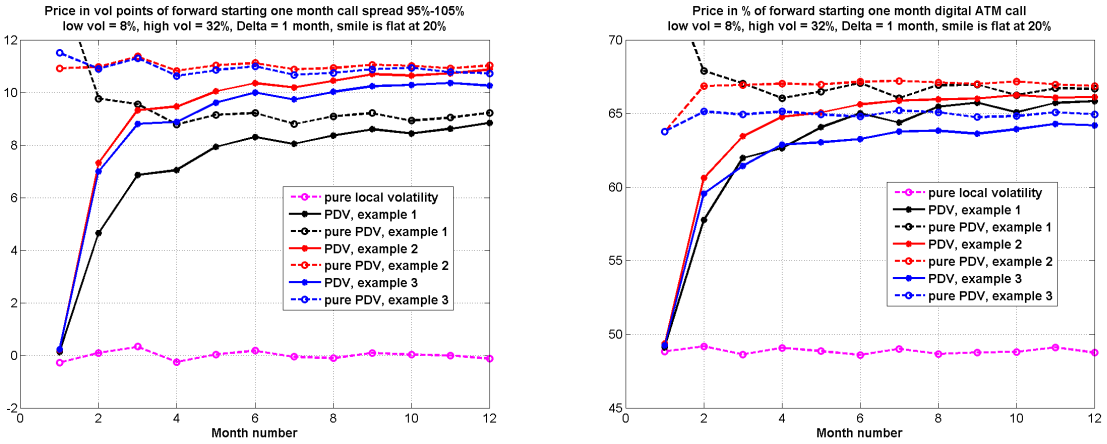

FIGURE 4.1. Prices in volatility points of forward starting call spreads $H_i^{\mathrm{CS}} \equiv (S_{T_i}/S_{T_{i-1}} - K_1)^+ - (S_{T_i}/S_{T_{i-1}} - K_2)^+$ (left) and prices in percent of forward starting digital calls $1\{S_{T_i}/S_{T_{i-1}} > 1\}$ (right) for LV, the PDV examples 1–3 of Table 1 (with their leverage function l), and their pure PDV counterparts; $T_i - T_{i-1} = 1$ month; $K_1 = 95\%$, $K_2 = 105\%$. On the x-axis, i is the month number. $\underline{\sigma} = 8\%$, $\overline{\sigma} = 32\%$, $\Delta = 1$ month. The LV and PDV models are calibrated to a flat smile (20%). For each model, the price in volatility points of the forward starting call spread is 100 times the value of $\Delta\sigma$ such that $\mathbb{E}^{\mathbb{Q}}[H_i^{\mathrm{CS}}] = C_{\mathrm{BS}}(K_1, \sigma_i^{\mathrm{ATM}} + \frac{\Delta\sigma}{2}) - C_{\mathrm{BS}}(K_2, \sigma_i^{\mathrm{ATM}} - \frac{\Delta\sigma}{2})$, where σ_i^{ATM} is the forward starting one-month ATM implied volatility in the corresponding model, i.e., $\mathbb{E}^{\mathbb{Q}}[(S_{T_i}/S_{T_{i-1}} - 1)^+] = C_{\mathrm{BS}}(1, \sigma_i^{\mathrm{ATM}})$; $C_{\mathrm{BS}}(K, \sigma)$ denotes the Black-Scholes price of a call with $S_0 = 1$, $r = q = 0$, and $T = 1$ month.

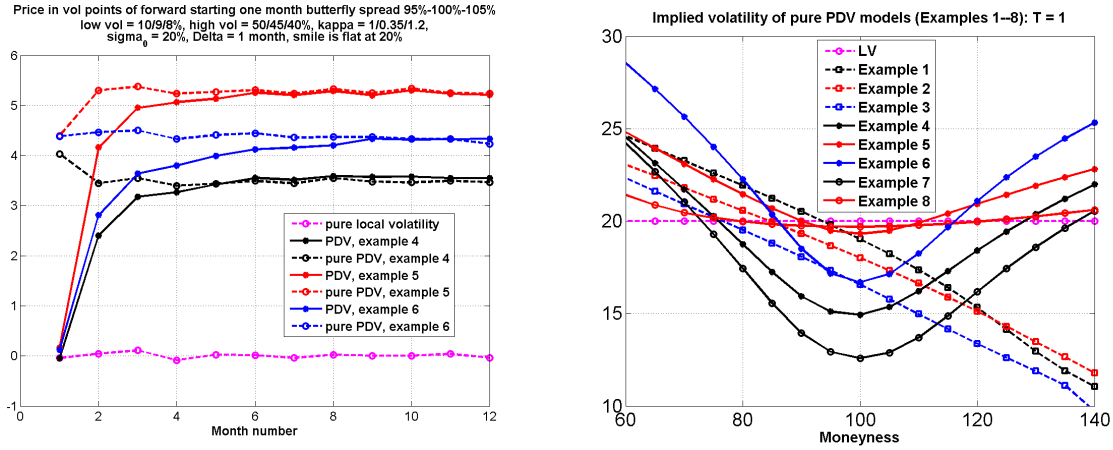

FIGURE 4.2. Left: Prices in volatility points of forward starting butterfly spreads $H_i^{\mathrm{B}} \equiv (S_{T_i}/S_{T_{i-1}} - K_1)^+ + (S_{T_i}/S_{T_{i-1}} - K_3)^+ - 2(S_{T_i}/S_{T_{i-1}} - K_2)^+$ for LV, the PDV examples 4–6 of Table 1 (with their leverage function l), and their pure PDV counterparts; $T_i - T_{i-1} = 1$ month; $K_1 = 95\%$, $K_2 = 100\%$, $K_3 = 105\%$. On the x-axis, i is the month number. $\underline{\sigma} = 10\%$ (resp. 9%, 8%), $\overline{\sigma} = 50\%$ (resp. 45%, 40%), $\kappa = 1$ (resp. 0.35, 1.2) for Example 4 (resp. 5, 6); $\sigma_0 = 20\%$, $\Delta = 1$ month. The LV and PDV models are calibrated to a flat smile (20%). For each example, the values of $\underline{\sigma}$ and $\overline{\sigma}$ are chosen so that the smile of the pure PDV model takes values around this flat smile (see right graph). For each model, the price in volatility points of the forward starting butterfly spread is 100 times the value of $\Delta\sigma$ such that $\mathbb{E}^{\mathbb{Q}}[H_i^{\mathrm{B}}] = C_{\mathrm{BS}}(K_1, \sigma_i^{\mathrm{ATM}} + \Delta\sigma) + C_{\mathrm{BS}}(K_3, \sigma_i^{\mathrm{ATM}} + \Delta\sigma) - 2C_{\mathrm{BS}}(K_2, \sigma_i^{\mathrm{ATM}})$; σ_i^{ATM} is the forward starting one-month ATM implied volatility in the corresponding model, i.e., $\mathbb{E}^{\mathbb{Q}}[(S_{T_i}/S_{T_{i-1}} - 1)^+] = C_{\mathrm{BS}}(1, \sigma_i^{\mathrm{ATM}})$.

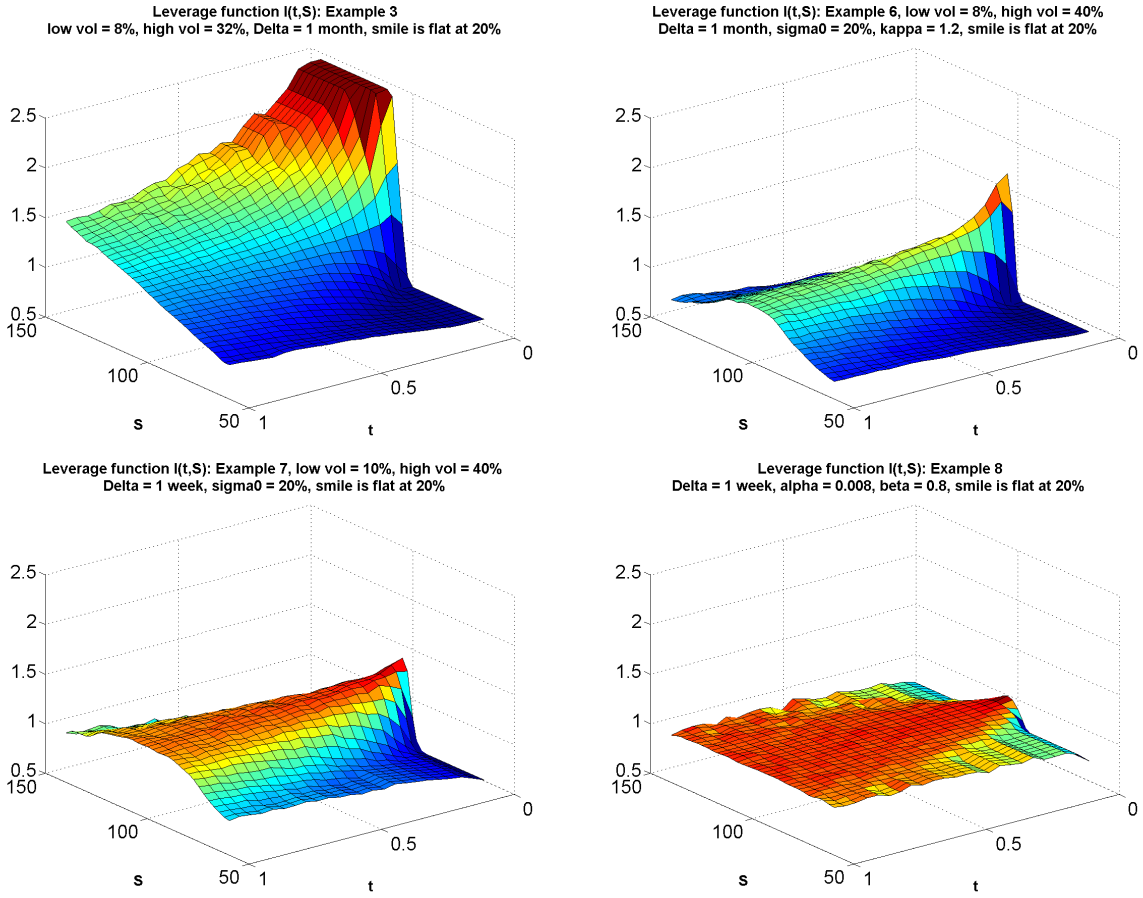

FIGURE 4.3. Leverage function $l(t, S)$ in the PDV models of Examples 3 (top left), 6 (top right), 7 (bottom left) and 8 (bottom right). Same parameter values as in Figure 4.2 for Examples 3 and 6. For Example 7: $\Delta = 1$ week, $\underline{\sigma} = 10\%$, $\overline{\sigma} = 40\%$, $\sigma_0 = 20\%$. For Example 8: $\Delta = 1$ week, $\alpha = 0.008$, $\beta = 0.8$. The models are calibrated to a flat smile (20%).

So far, we have shown that PDV models can do as well as SLV models—and even better than SLV models, considering the unreasonably high levels of the corresponding vol of vol—to reproduce popular properties of the smile dynamics, while staying calibrated to a smile, with the advantage that they are complete and more parsimonious. But PDV models can do much more than that: they have so many degrees of freedom—the path-dependent variables X, and the function $\sigma(t, S, X)$—that they can generate spot-vol dynamics that are *not* attainable by SLV models. For example, imagine that a sophisticated client—a hedge fund at the cutting hedge of innovation for instance—asks a quote on the conditional variance swap with payoff

$$H_T = \sum_{i=1}^{n-1} r_{i+1}^2 1_{\{r_i \le 0\}} \approx \int_0^T \sigma_t^2 1_{\left\{\frac{S_t}{S_{t-\Delta}} \le 1\right\}} \, dt, \qquad r_i = \frac{S_{t_i} - S_{t_{i-1}}}{S_{t_{i-1}}}, \qquad \Delta = t_i - t_{i-1} = 1\,\text{day} \tag{4.1}$$

Then, in the SLV model, for a given risk-neutral probability $\mathbb{Q}$,

$$\mathbb{E}^{\mathbb{Q}}\left[\sigma_t^2 1_{\left\{\frac{S_t}{S_{t-\Delta}} \le 1\right\}} \middle| S_t\right] \approx \mathbb{E}^{\mathbb{Q}}\left[\sigma_t^2 \middle| S_t\right] \mathbb{E}^{\mathbb{Q}}\left[1_{\left\{\frac{S_t}{S_{t-\Delta}} \le 1\right\}} \middle| S_t\right] \approx \frac{1}{2}\sigma_{\text{Dup}}^2(t, S_t)$$

since, given the asset price, the volatility level is almost independent of the sign of the past daily asset return, and the SLV model is calibrated to the market smile. Thus both the SLV price and the LV price are very close to the variance swap price halved:

$$\text{SLV price} \;\approx\; \text{LV price} \;\approx\; \frac{1}{2}\int_0^T \mathbb{E}^{\mathbb{Q}}\left[\sigma_{\text{Dup}}^2(t, S_t)\right] dt \;=\; \frac{1}{2}\text{var swap price}$$

whatever the choice of $\mathbb{Q}$. However, if a hedge fund requests such a quote, it is probably because they observed that the historical time series of this asset price shows that r_{i+1}^2 is large when $r_i \leq 0$, and small otherwise, expect this to continue, and try to statistically arbitrage a counterparty. *All the common models* that are used in the industry today would fail to capture this risk, but the PDV model of Example 1, with $\Delta = t_i - t_{i-1} = 1$ day, grasps it very well, since precisely $\sigma(t_i, S_{t_i}, S_{t_{i-1}})$ is high ($\overline{\sigma}$) when $r_i \leq 0$, and low ($\underline{\sigma}$) otherwise. As Δ is small, $\mathbb{E}^{\mathbb{Q}}\left[\sigma(t_i, S_{t_i}, S_{t_{i-1}})^2 \,|S_{t_i}\right] \approx \frac{\overline{\sigma}^2+\underline{\sigma}^2}{2}$ is almost constant, so

$$\text{PDV price} \approx \int_0^T \mathbb{E}^{\mathbb{Q}}\left[\frac{\overline{\sigma}^2}{\frac{\overline{\sigma}^2+\underline{\sigma}^2}{2}}\sigma_{\text{Dup}}^2(t,S_t)1_{\left\{\frac{S_t}{S_{t-\Delta}}\leq 1\right\}}\right] dt \approx \frac{\overline{\sigma}^2}{\overline{\sigma}^2+\underline{\sigma}^2}\,\text{var swap price}$$

which, for reasonable values of $(\underline{\sigma}, \overline{\sigma})$, e.g., $(10\%, 40\%)$, is close to the (unconditional) variance swap price, i.e., twice the SLV price and twice the LV price. In such a case, an investment bank equipped with PDV may avoid a large mispricing. Note that for the pure PDV model,

$$\text{pure PDV price} \approx \int_0^T \mathbb{E}^{\mathbb{Q}}\left[\overline{\sigma}^2 1_{\left\{\frac{S_t}{S_{t-\Delta}}\leq 1\right\}}\right] dt \approx \frac{\overline{\sigma}^2 T}{2} \approx \frac{\overline{\sigma}^2}{\overline{\sigma}^2+\underline{\sigma}^2}\,\text{pure PDV var swap price}$$

5. Choose a particular PDV to capture historical patterns of volatility

With so many degrees of freedom, PDV models can easily capture historical properties of volatility. This means that, like the local correlation models presented in [11], PDV models are flexible enough to reconcile implied calibration (e.g., calibration to the market smile) with historical calibration (calibration from historical time series of asset prices): one chooses a PDV $\sigma(t, S, X)$ from the observation of the time series, e.g., one observed that the short term ATM volatility is a certain function of $S_t/\overline{S}_t^{\Delta}$, then multiplies it by a leverage function $l(t, S)$ and eventually calibrates l to the market smile using the particle method. Often, the leverage function flattens over time so the calibrated model still has the desired property, except maybe close to $t = 0$. By construction, PDV models are flexible enough to capture *any* path-dependency of the volatility. For a given choice of PDV, what remains to be numerically checked is how much and how long the smile calibration distorts the link between past prices and current instantaneous volatility, and whether the model produces reasonable dynamics of implied volatility.

For instance, Examples 1–3 above relate the *level* of volatility to recent *relative changes (returns)* in the asset price, allowing volatility to rise *very quickly* (from, say, 10% to 20%, a 100% return) in periods when the asset price goes down, regardless of the price level—a pattern which is obvious on S&P 500 data (see Figure 5.1, left). Actually, the two basic quantities that possess a natural scale are the *level* of volatility and the *relative changes* in the asset price so we believe that a good model should relate these two quantities. In this respect, the LV model, which links the *level* of volatility to the *level* of the asset price, does not make much financial sense—so well designed PDV models need not be recalibrated as often as the LV model. SV models, which connect the *change* in volatility to the *relative change* in the asset price, do not make much more sense: only unreasonable levels of vol of vol allow large movements (e.g., a 100% return) of instantaneous volatility, and therefore a large mean-reversion is artificially added to keep volatility within its natural range. By contrast PDV models can easily capture such large changes in volatility. Examples 1–6 are extreme cases, where the instantaneous volatility jumps from $\underline{\sigma}$ to $\overline{\sigma}$ or conversely; in practice smoothed versions may make more financial sense.

Another obvious pattern visible on Figure 5.1 (left) is the boundedness of volatility paths: the one-month ATM implied volatility path stays away from zero and does not take extreme values. This is easily enforced in PDV models, but is not always the case for popular SV models such as the Heston model, in which a large proportion of volatility paths get close to zero, while others may take very high values, if the Feller condition is not satisfied. Enforcing bounded volatility also prevents potential problems of loss of asset martingality.

The dependency assumptions of Examples 1–3 can be tested on historical data. From Figure 5.1 (left), it is clear that for the S&P 500, the level of volatility is *not* determined by the level of the asset price, but by the recent changes in the asset price: spikes in the one-month ATM implied volatility correspond to those periods where the asset price is below its moving average, creating volatility clustering.

To reproduce Figure 5.1 (left), we used the actual S&P 500 time series and generated the one-month ATM implied volatility path using the pure PDV model of Example 2, with $\underline{\sigma} = 8\%$, $\overline{\sigma} = 21\%$, and $\Delta = 1$ month (see Figure 5.1, right). The similarity with the actual volatility path is striking. Even though the

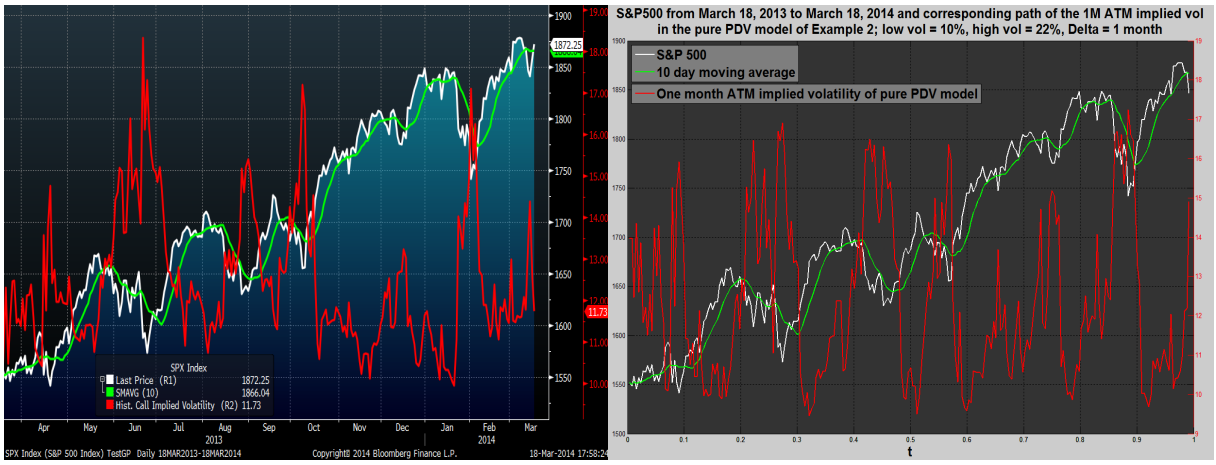

FIGURE 5.1. Left: Historical time series of the S&P 500 (white), its 10 day moving average (green), and a short-dated ATM implied volatility ($T \approx 1$ month), from March 18, 2013 to March 18, 2014. Copyright 2014 Bloomberg Finance L.P. Right: the one-month ATM implied volatility is built using the actual time series of the S&P 500 (white) and the pure PDV model of Example 2, with $\underline{\sigma} = 8\%$, $\overline{\sigma} = 21\%$, and $\Delta = 1$ month.

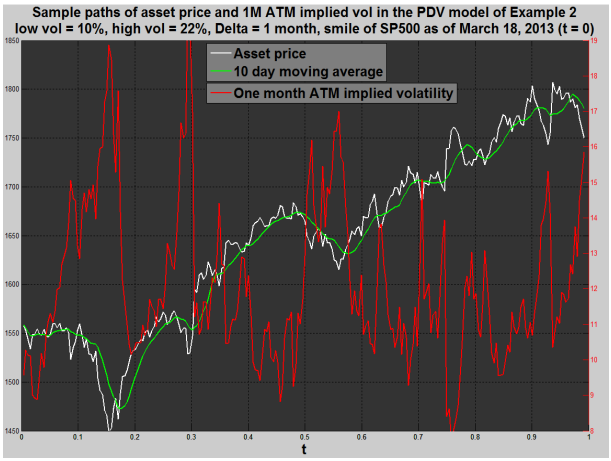

FIGURE 5.2. Sample paths of the S&P 500 (white), its 10 day moving average (green), and the one-month ATM implied volatility using the PDV model of Example 2, calibrated to the smile of the S&P 500 on March 18, 2013, with $\underline{\sigma} = 10\%$, $\overline{\sigma} = 22\%$, and $\Delta = 1$ month.

(unobservable) instantaneous volatility takes only two values, 8% and 21%, the one-month ATM implied volatility, which is an average of future instantaneous volatilities, varies continuously, with spikes when the market is locally bearish. In Figure 5.2 we pick $\underline{\sigma} = 10\%$ and $\overline{\sigma} = 22\%$ and calibrate the leverage function to the smile of the S&P 500 as of March 18, 2013, the first day of the one-year time window of Figure 5.1 (left). The implied and local volatility surfaces are reported in Figure 5.3. Calibrating to the smile hardly affects the spot-vol dynamics: the leverage function (Figure 5.4, left) distorts only slightly the causal link between asset returns and volatility, mostly just after the calibration date. In contrast with the pure PDV model, it also makes it possible for the implied volatility to lie below $\underline{\sigma}$ or above $\overline{\sigma}$. Note that the pure PDV smile is flatter than the market smile (Figure 5.4, right): part of the smile is produced by the pure PDV, and the rest by the leverage function l. On this data, the shape of l is inverted at very short maturities, for which the ATM market skew is smaller (in absolute value) than the ATM skew of the pure PDV model.

With such parameter values, the PDV model of Example 2 captures what we believe is a major pattern of the historical joint behaviour of the S&P 500 and its short term implied volatilities. But what about pricing? Here, the volatility interval $[\underline{\sigma}, \overline{\sigma}]$ is not as wide as in Figure 4.1, so we expect forward starting call spreads

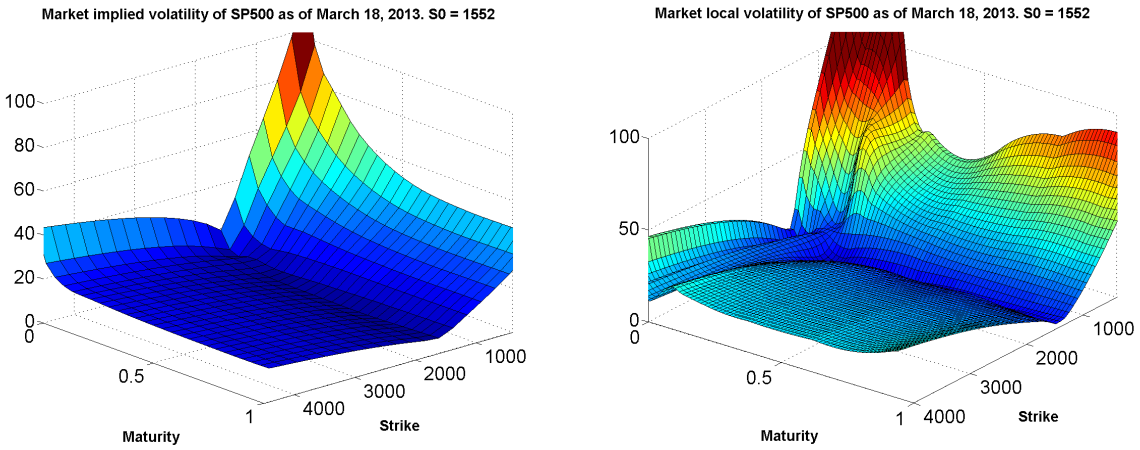

FIGURE 5.3. Surfaces of implied volatility (left) and local volatility (right) of the S&P 500 as of March 18, 2013.

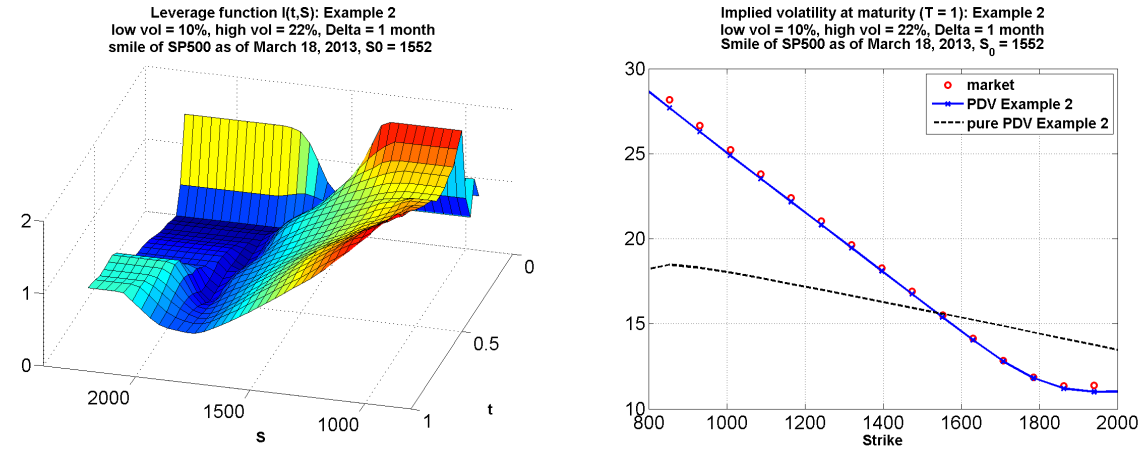

FIGURE 5.4. Left: leverage function $l(t, S)$ for the PDV model of Example 2, with $\underline{\sigma} = 10\%$, $\overline{\sigma} = 22\%$, and $\Delta = 1$ month, calibrated to the smile of the S&P 500 as of March 18, 2013. Right: Market, PDV, and pure PDV implied volatilities at maturity ($T = 1$). $S_0 = 1552$.

and digital calls to be cheaper. It is indeed the case. However, the PDV forward skew is still sizeably larger (around 5 volatility points larger, for the 95%-105% skew) than the LV forward skew (see Figure 5.5).

The fact that volatility depends on recent asset returns is also supported by other statistical analyses [19, 7]. Other empirical studies show that the volatility may depend on recent realized volatility. So far, only the ARCH (Engle, 1982) and GARCH (Bollersev, 1986) models, which are popular among econometricians, could capture this. These models are particularly appreciated for capturing tail heaviness, volatility clustering and dependence without correlation, like Examples 1–6 above. Our approach generalizes them by defining *local ARCH models*, in which (a) the ARCH volatility is multiplied by a leverage function in order to fit a smile, and (b) the function $\sigma(X)$ is arbitrary:

$$\frac{dS_t}{S_t} = \sigma(X_t) l(t, S_t)\, dW_t, \qquad X_t = \sum_{t-\Delta < t_i \le t} r_i^2, \qquad r_i = \frac{S_{t_i} - S_{t_{i-1}}}{S_{t_{i-1}}}$$

For example, if we choose $\sigma(t, X) = \underline{\sigma}$ if $X \le \sigma_0$ and $\overline{\sigma}$ otherwise for a given σ_0 (Example 7), then we get the leverage function $l(t, S)$ of Figure 4.3 (bottom left), vanishing ATM forward skew (because the inputed smile is flat and the ARCH volatility depends only on *squared* returns), and values of forward starting butterfly spreads around 2.4 points of volatility (to be compared with the values in Figure 4.2). To mimic ARCH

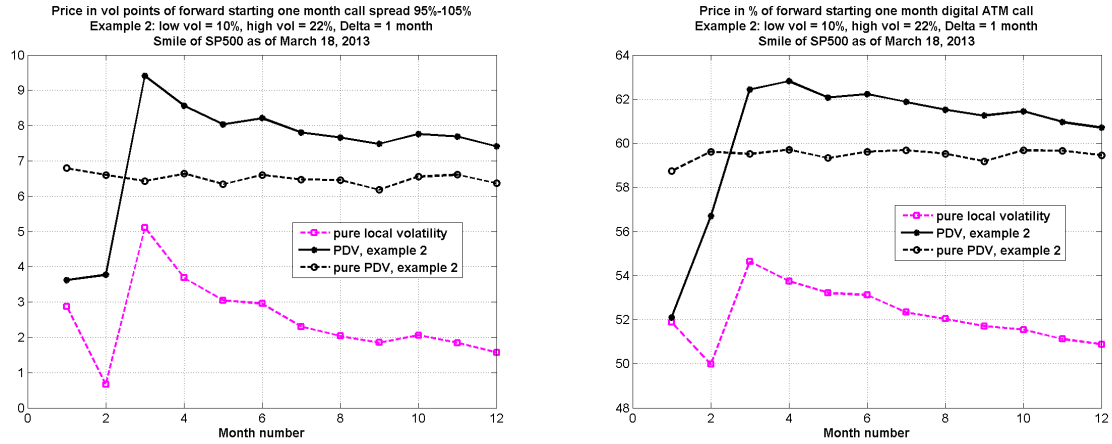

FIGURE 5.5. Prices in volatility points of forward starting call spreads $H_i^{\mathrm{CS}} \equiv (S_{T_i}/S_{T_{i-1}} - K_1)^+ - (S_{T_i}/S_{T_{i-1}} - K_2)^+$ (left) and prices in percent of forward starting digital calls $1\{S_{T_i}/S_{T_{i-1}} > 1\}$ (right) for LV, the PDV example 2 of Table 1 (with its leverage function l), and its pure PDV counterpart; $T_i - T_{i-1} = 1$ month; $K_1 = 95\%$, $K_2 = 105\%$. On the x-axis, i is the month number. $\underline{\sigma} = 10\%$, $\overline{\sigma} = 22\%$, $\Delta = 1$ month. The LV and PDV models are calibrated to the smile of the S&P 500 as of March 18, 2013. For each model, the price in volatility points of the forward starting call spread is 100 times the value of $\Delta\sigma$ such that $\mathbb{E}^{\mathbb{Q}}[H_i^{\mathrm{CS}}] = C_{\mathrm{BS}}(K_1, \sigma_i^{\mathrm{ATM}} + \frac{\Delta\sigma}{2}) - C_{\mathrm{BS}}(K_2, \sigma_i^{\mathrm{ATM}} - \frac{\Delta\sigma}{2})$; σ_i^{ATM} is the forward starting one-month ATM implied volatility in the corresponding model, i.e., $\mathbb{E}^{\mathbb{Q}}[(S_{T_i}/S_{T_{i-1}} - 1)^+] = C_{\mathrm{BS}}(1, \sigma_i^{\mathrm{ATM}})$.

models, we can choose $\sigma(X)^2 = \alpha + \beta X$ with $\alpha > 0$, $\beta < 1$ (Example 8); we then get a much flatter pure PDV smile and a much flatter leverage function l (see Figures 4.2 and 4.3, bottom right), vanishing ATM forward skew, and values of forward starting butterfly spreads around 0.7 point of volatility. Obviously if needed one can weight the squared returns in the above definition of X. To generalize GARCH processes, one can include past values of $\sigma(X_{t_i})$ in the definition of X. For a modification of GARCH processes that captures assymetry in returns and a suggestion of a continuous-time limit, see [14]. A continuous-time version of the GARCH model, driven by a Lévy process, was proposed by Klüppelberg, Lindner and Maller [18].

6. GENERALIZATIONS

It is easy to generalize to path-dependent interest rates and dividend yield. From Proposition 12.8 in [10] (page 377), the (complete) model

$$\frac{dS_t}{S_t} = (r_t - q_t)\, dt + \sigma_t l(t, S_t)\, dW_t$$

where $r_t \equiv r(t, (S_u, u \le t))$, $q_t \equiv q(t, (S_u, u \le t))$ and $\sigma_t \equiv \sigma(t, (S_u, u \le t))$ are path-dependent, is calibrated to the market smile of S if and only if for all (t, K)

(6.1)

$$l(t,K)^2 \frac{\mathbb{E}[D_{0t}\sigma_t^2 | S_t = K]}{\mathbb{E}[D_{0t} | S_t = K]} = \sigma_{\mathrm{Dup}}(t,K)^2 - \frac{\mathbb{E}\left[D_{0t}\left(r_t - q_t - (r_t^0 - q_t^0)\right) 1_{S_t > K}\right]}{\frac{1}{2}K\partial_K^2 \mathcal{C}(t,K)} + \frac{\mathbb{E}\left[D_{0t}\left(q_t - q_t^0\right)(S_t - K)^+\right]}{\frac{1}{2}K^2\partial_K^2 \mathcal{C}(t,K)}$$

where $D_{0t} = \exp\left(-\int_0^t r_s ds\right)$ is the discount factor, r_t^0 and q_t^0 are deterministic rates and repos, and

$$\sigma_{\mathrm{Dup}}(t,K)^2 = \frac{\partial_t \mathcal{C}(t,K) + (r_t^0 - q_t^0) K \partial_K \mathcal{C}(t,K) + q_t^0 \mathcal{C}(t,K)}{\frac{1}{2}K^2\partial_K^2 \mathcal{C}(t,K)}$$

is the Dupire LV, with $\mathcal{C}(t,K)$ the market price of the call option on S with strike K and maturity t. Hence the particle method can be used to calibrate l. Condition (6.1) is actually valid even if r_t, q_t and σ_t are

also driven by Brownian motions other than W, allowing us to define stochastic path-dependent volatility (SPDV) models that are calibrated to the market smile—but then we loose completeness.

7. Conclusion

In this article we have shown that path-dependent volatility models are excellent candidates to challenge the duopoly of local volatility and stochastic volatility which has dominated option pricing over the last 20 years. They indeed combine benefits from both LV and SV: like the LV model they are complete and can fit the market smile, so all derivatives have a unique price consistent with today's prices of vanilla options; and like SV models they can produce rich spot-vol dynamics, such as large negative short term forward skews, large forward smile curvatures, or large vols of vol—but in a more parsimonious way, since no extra Brownian motion is needed. Thanks to their huge flexibility—one may choose any set of path-dependent variables X and any PDV $\sigma(t, S, X)$—PDV models actually span a much broader range of spot-vol dynamics than SV models, possibly preventing large mispricings, and they can also capture important historical features of asset prices and volatilities, such as volatility *levels* depending on recent asset *returns*, tail heaviness, volatility clustering, and dependence without correlation.

In practice, the particle method is so simple and efficient that the smile calibration is not a problem: efforts can be concentrated on the choice of a convenient PDV, depending on the market and derivative under consideration. To this end, we hope that the examples we showcased will serve as a helpful guide. Beyond the ability to produce desired spot-vol dynamics and capture spot-vol historical patterns, an important criterion to assess the quality of a PDV model should be its hedging performance on backtests, a task we leave for future work.

Acknowledgements. I am grateful to Lorenzo Bergomi, Sylvain Corlay, Stéphane Crapanzano, Bruno Dupire, Pierre Henry-Labordère, Bryan Liang, and two anonymous referees for their valuable feedback. I also thank one of the two referees for pointing out reference [3] to me, and Paul Embrechts for reference [18].

References

[1] Bergomi, L.: *Smile dynamics II*, Risk, October, 2005.
[2] Bollerslev, T.: *Generalized autoregressive conditional heteroskedasticity*, Journal of Econometrics, 31:307-327, 1986.
[3] Brunick, G. and Shreve, S.: *Mimicking an Itô process by a solution of a stochastic differential equation*, Ann. Appl. Prob., 23(4):1584–1628, 2013.
[4] Dupire, B.: *Pricing with a smile*, Risk, January, 1994.
[5] Engle, R.: *Autoregressive conditional heteroscedasticity with estimates of variance of United Kingdom inflation*, Econometrica 50:987-1008, 1982.
[6] Figà-Talamanca, G. and Guerra, M.L.: *Fitting prices with a complete model*, J. Bank. Finance 30(1), 247–258, 2006.
[7] Foschi, P. and Pascucci, A.: *Path dependent volatility*, Decisions Econ. Finan., 2007.
[8] Fouque, J.-P., Papanicolaou, G. and Sircar, R.: *Derivatives in financial markets with stochastic volatility*, Cambridge University Press, 2000.
[9] Guyon, J. and Henry-Labordère, P.: *Being particular about calibration*, Risk, January, 2012.
[10] Guyon, J. and Henry-Labordère, P.: *Nonlinear option pricing*, Chapman & Hall/CRC Financial Mathematics Series, 2013.
[11] Guyon, J.: *Local correlation families*, Risk, February, 2014.
[12] Henry-Labordère, P.: *Dupire-like formulas for path-dependent options*, private communication.
[13] Heston, S.: *A closed-form solution for options with stochastic volatility with applications to bond and currency options*, The Review of Financial Studies, 6(2):327–343, 1993.
[14] Heston, S. and Nandi, S.: *Preference-free option pricing with path-dependent volatility: a closed-form approach*, working paper from Federal Reserve Bank of Atlanta, 1998.
[15] Hobson, D. G. and Rogers, L.C.G.: *Complete models with stochastic volatility*, Mathematical Finance 8 (1), 27–48, 1998.
[16] Hubalek, F., Teichmann, J. and Tompkins, R.: *Flexible complete models with stochastic volatility generalising Hobson-Rogers*, working paper, 2004.
[17] Hull, J. and White, A.: *The pricing of options with stochastic volatilities*, The Journal of Finance, 42(2):281–300, 1987.
[18] Klüppelberg, C., Lindner, A. and Maller, R.: *A continuous time GARCH process driven by a Lévy process: stationarity and second order behaviour*, J. Appl. Probab., 41(3):601–622, 2004.
[19] Platania, A. and Rogers, L.C.G.: *Putting the Hobson-Rogers model to the test*, working paper, 2003.

Quantitative Research, Bloomberg L.P., 731 Lexington Avenue, New York, NY 10022, USA.