On the Guyon–Lekeufack Volatility Model[*]

Marcel Nutz[†]   Andrés Riveros Valdevenito[‡]

July 3, 2023

Abstract

Guyon and Lekeufack recently proposed a path-dependent volatility model and documented its excellent performance in fitting market data and capturing stylized facts. The instantaneous volatility is modeled as a linear combination of two processes, one is an integral of weighted past price returns and the other is the square-root of an integral of weighted past squared volatility. Each of the weightings is built using two exponential kernels reflecting long and short memory. Mathematically, the model is a coupled system of four stochastic differential equations. Our main result is the wellposedness of this system: the model has a unique strong (non-explosive) solution for realistic parameter values. We also study the positivity of the resulting volatility process.

Keywords Path-dependent volatility model, SDE, wellposedness, explosion
AMS 2010 Subject Classification 60H10; 91G20

1 Introduction

Path-dependent volatility models (PDV) are stochastic models for security prices where the instantaneous volatility is a function of the price path. Starting with [5, 2, 9, 1, 3], such models emphasize that prices have a feedback on volatility (e.g., the leverage effect) rather than the volatility being an exogenous factor driving the price. In their recent paper [4], Guyon and Lekeufack empirically study the volatility of the S&P 500 index (and other indexes) and conclude that the majority of the variation can be explained by past index returns. Indeed, the relevant statistics are 1. weighted sum of past daily returns and 2. square-root of weighted sum of past daily squared returns (i.e., squared volatility). More specifically, long and short memory are both found to be important, hence the authors recommend using two decay kernels with different time scales. This leads to four processes feeding into the volatility: weighted sum of past returns at two timescales (indexed as (1,0) and (1,1) below) and weighted sum of past squared returns at two (different) timescales (indexed as (2,0) and (2,1) below).

For practical purposes, [4] finds that exponential kernels provide a tractable model with good fit. This leads the authors to propose a Markovian model with nine parameters, called the Markovian 4-factor PDV model. They convincingly argue that this model captures the important stylized facts of volatility, produces realistic price and volatility paths, and can

jointly fit S&P 500 and VIX smiles. Specifically, the volatility process of the 4-factor PDV model is given as

$$\sigma_t = \beta_0 + \beta_1 R_{1,t} + \beta_2 \sqrt{R_{2,t}}.$$

Here $R_{1,t}$ is the convex combination $(1-\theta_1)R_{1,0,t} + \theta_1 R_{1,1,t}$ of the past returns weighted with different decay rates $\lambda_{1,j}$; i.e., $R_{1,j,t}$ is an Ornstein–Uhlenbeck process $dR_{1,j,t} = \lambda_{1,j}\sigma_t dW_t - \lambda_{1,j}R_{1,j,t}dt$ for $j \in \{0,1\}$. Moreover, $R_{2,t}$ is a convex combination $(1-\theta_2)R_{2,0,t} + \theta_2 R_{2,1,t}$ of the past squared volatility weighted with different decay rates $\lambda_{2,j}$; i.e., $R_{2,j,t}$ is an exponential moving average $dR_{2,j,t} = \lambda_{2,j}\left(\sigma_t^2 - R_{2,j,t}\right)dt$ for $j \in \{0,1\}$.

Altogether, this leads to a coupled SDE system for the four processes $(R_{i,j,t})$, stated as (4-PDV) below. Due to the square and square-root terms in the dynamics, its wellposedness is not obvious. Most importantly, it is not clear if the system explodes in finite time (the numerical simulations in [4] truncate the volatility at a fixed upper bound). The purpose of this paper is to provide existence and uniqueness for (4-PDV) under conditions compatible with the parameter values required in practice.

The main results are summarized in the subsequent section. There, we first discuss a simpler model which uses only one timescale for each process $R_{i,t}$, corresponding to the special case $\theta_i \in \{0,1\}$. While [4] details that this 2-factor model does not provide a good fit in practice, the authors find it useful to gain intuition about the more complicated 4-factor model. Following their didactic lead, we first prove our results for the 2-factor model in Section 2. In this case, the equations are simpler and the algebraic expressions clearly motivate our strategy of proof. Guided by those insights, the 4-factor model can be treated using a similar strategy (detailed in Section 3), though the expressions are more convoluted.

1.1   Main Results

The 2-factor model of [4] is specified by an SDE driven by a standard Brownian motion W,

$$\begin{aligned}
\sigma_t &= \beta_0 + \beta_1 R_{1,t} + \beta_2 \sqrt{R_{2,t}} \\
dR_{1,t} &= \lambda_1 \sigma_t dW_t - \lambda_1 R_{1,t} dt \\
dR_{2,t} &= \left(\lambda_2 \sigma_t^2 - \lambda_2 R_{2,t}\right) dt
\end{aligned} \tag{2-PDV}$$

with parameters

$$\beta_0, \beta_2, \lambda_1, \lambda_2 \geq 0 \quad \text{and} \quad \beta_1 \leq 0$$

and initial values $R_{1,0} \in \mathbb{R}$ and $R_{2,0} \in (0,\infty)$. The above is an autonomous SDE for the processes $(R_{1,t}, R_{2,t})$, with σ_t merely acting as an abbreviation. On the other hand, if σ_t is given, the equations for $R_{1,t}$ and $R_{2,t}$ in (2-PDV) are straightforward: $R_{1,t}$ is an Ornstein–Uhlenbeck process driven by the log-returns $\sigma_t dW_t$,

$$R_{1,t} = R_{1,0}e^{-\lambda_1 t} + \lambda_1 \int_0^t e^{-\lambda_2(t-s)}\sigma_s dW_s,$$

and $R_{2,t}$ is an exponential moving average of σ_t^2,

$$R_{2,t} = R_{2,0}e^{-\lambda_2 t} + \lambda_2 \int_0^t \sigma_s^2 e^{-\lambda_2(t-s)}ds > 0. \tag{1}$$

In particular, the expression $\sqrt{R_{2,t}}$ in (2-PDV) is well-defined.

2

Theorem 1. *Suppose that $\beta_2 < 1$ and $\lambda_1 \beta_1^2 < 2$. Then the 2-factor model* (2-PDV) *has a unique strong solution.*

The proof is detailed in Section 2. There, we observe that strong existence and uniqueness always hold up to a possible explosion time; the parameter restrictions in Theorem 1 are used to prove absence of explosions in finite time. Table 1 reports the parameters used in [4]. We see that the parameter restrictions in Theorem 1 are easily satisfied for those values. In Theorem 7 of Section 2, we also report the condition $\lambda_2 < 2\lambda_1$ ensuring $\sigma_t > 0$; that condition is also satisfied by the values in Table 1.

β_0	β_1	β_2	λ_1	λ_2
0.08	-0.08	0.5	62	40

Table 1: Example parameters for the 2-factor model from [4, Table 7]

Next, we move on to the 4-factor model of [4]. It is specified by the SDE

$$
\begin{aligned}
\sigma_t &= \beta_0 + \beta_1 R_{1,t} + \beta_2 \sqrt{R_{2,t}} \\
R_{1,t} &= (1 - \theta_1) R_{1,0,t} + \theta_1 R_{1,1,t} \\
R_{2,t} &= (1 - \theta_2) R_{2,0,t} + \theta_2 R_{2,1,t} \\
dR_{1,j,t} &= \lambda_{1,j} \sigma_t dW_t - \lambda_{1,j} R_{1,j,t} dt, \quad j \in \{0,1\} \\
dR_{2,j,t} &= \lambda_{2,j} \left(\sigma_t^2 - R_{2,j,t} \right) dt, \quad j \in \{0,1\}
\end{aligned}
\tag{4-PDV}
$$

with parameters

$$
\beta_0, \beta_2, \lambda_{1,j}, \lambda_{2,j} \geq 0, \qquad \beta_1 \leq 0, \qquad \theta_1, \theta_2 \in [0,1]
$$

and initial values $R_{1,j,0} \in \mathbb{R}$ and $R_{2,j,0} > 0$, $j \in \{0,1\}$. The above is an autonomous SDE for the four processes $(R_{1,j,t}, R_{2,j,t})_{j \in \{0,1\}}$. We note that (4-PDV) generalizes the 2-factor model (2-PDV); the latter is recovered when $\theta_1, \theta_2 \in \{0,1\}$. Whereas for $\theta_1, \theta_2 \in (0,1)$, the difference with (2-PDV) is that $R_{1,t}, R_{2,t}$ are proper convex combinations of processes with different time scales. Once again, $R_{1,j,t}$ and $R_{2,j,t}$ have straightforward expression once σ_t is given, and $R_{2,j,t} > 0$ as in (1). Our main result reads as follows.

Theorem 2. *Suppose that*

$$
\beta_2^2 \theta_2 < 1, \qquad \beta_2^2 (1 - \theta_2) < 1, \qquad \lambda_{1,0} \beta_1^2 (1 - \theta_1)^2 < 2, \qquad \lambda_{1,1} \beta_1^2 \theta_1^2 < 2.
$$

Then the 4-factor model (4-PDV) *has a unique strong solution.*

The proof is stated in Section 3. Again, strong existence and uniqueness always hold up to a possible explosion time. On the other hand, the example parameters determined in [4], reproduced in Table 2, comfortably satisfy the parameter restrictions in Theorem 2. In conclusion, explosions do not happen for realistic parameters. Finally, we observe that the restrictions in Theorem 2 reduce to the ones in Theorem 1 when $\theta_1, \theta_2 \in \{0,1\}$, so that Theorem 2 fully contains Theorem 1.

3

β_0	β_1	β_2	$\lambda_{1,0}$	$\lambda_{1,1}$	$\lambda_{2,0}$	$\lambda_{2,1}$	θ_1	θ_2
0.04	-0.13	0.65	55	10	20	3	0.25	0.5

Table 2: Example parameters for the 4-factor model from [4, Table 8]

2 Analysis of the 2-Factor Model (2-PDV)

We first show, using fairly standard arguments, that (2-PDV) has a unique strong solution up to a possible explosion time. In the second part of this section, we prove the absence of explosions under the parameter restrictions stated in Theorem 1. In the final part, we study the positivity of σ_t.

To detail the aforementioned arguments, we introduce a more concise notation for (2-PDV): writing $R_t := (R_{1,t}, R_{2,t})$, we can rewrite (2-PDV) as

$$dR_t = b(R_t)dt + \nu(R_t)dW_t, \qquad R_0 = (R_{1,0}, R_{2,0})$$

$$\nu(x, y) = \begin{pmatrix} \lambda_1 (\beta_0 + \beta_1 x + \beta_2 \sqrt{y}) \\ 0 \end{pmatrix}$$

$$b(x, y) = \begin{pmatrix} -\lambda_1 x \\ \lambda_2 \left(\beta_0^2 + \beta_1^2 x^2 + (\beta_2^2 - 1)y + 2\beta_0\beta_1 x + 2\beta_0\beta_2 \sqrt{y} + 2\beta_1\beta_2 x \sqrt{y} \right) \end{pmatrix}.$$

As the coefficients $\nu(x, y)$ and $b(x, y)$ are continuous in their domains and the initial condition is deterministic, the general existence result of [6, Theorem IV.2.3] shows the following.

Lemma 3. *The SDE* (2-PDV) *has a weak solution up to a possible explosion time.*

Next, we establish pathwise uniqueness. The usual local Lipschitz condition (e.g., [6, Theorem IV.3.1]) fails because of the term $\sqrt{y}$ in the coefficients. However, as this failure only occurs at the boundary of the relevant domain, a modification of the usual proof applies.

Lemma 4. *The SDE* (2-PDV) *satisfies pathwise uniqueness.*

Proof. Following the proof of [6, Theorem IV.3.1], we consider two solutions (R, W) and (R', W) of (2-PDV) on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with Brownian motion W, and with the same initial values $R_0 = R'_0 = (R_{1,0}, R_{2,0})$.

Given $N, \varepsilon > 0$, there exists $K_{\varepsilon, N} > 0$ such that

$$\left\| \nu(x, y) - \nu(x', y') \right\|^2 + \left\| b(x, y) - b(x', y') \right\|^2 \leq K_{\varepsilon, N} \left\| (x, y) - (x', y') \right\|^2$$

for all $(x, y), (x', y') \in \mathbb{R}^2$ with $\|(x, y)\|, \|(x', y')\| \leq N$ and $y, y' \geq \varepsilon$. Define

$$S_N := \inf \{t \geq 0 : \|R_t\| \geq N\}, \qquad T_\varepsilon := \inf \{t \geq 0 : |R_{2,t}| \leq \varepsilon\}$$

and similarly S'_N, T'_ε for R' instead of R. Set $S_{\varepsilon, N} := S_N \wedge S'_N \wedge T_\varepsilon \wedge T'_\varepsilon$ and note that

$$R_{t \wedge S_{\varepsilon, N}} - R'_{t \wedge S_{\varepsilon, N}} = \int_0^{t \wedge S_{\varepsilon, N}} \left(\nu(R_s) - \nu(R'_s) \right) dW_s + \int_0^{t \wedge S_{\varepsilon, N}} \left(b(R_s) - b(R'_s) \right) ds.$$

4

Fix $T \in (0, \infty)$. For $t \leq T$, Itô's isometry and Hölder's inequality yield

$$
\mathbb{E}\left(\| R_{t \wedge S_{\varepsilon,N}} - R'_{t \wedge S_{\varepsilon,N}} \|^2 \right)
$$
$$
\leq 2\mathbb{E}\left(\left| \int_0^{t \wedge S_{\varepsilon,N}} \left(\nu(R_s) - \nu(R'_s) \right) dW_s \right|^2 \right) + 2\mathbb{E}\left(\left| \int_0^{t \wedge S_{\varepsilon,N}} \left(b(R_s) - b(R'_s) \right) ds \right|^2 \right)
$$
$$
\leq 2\mathbb{E}\left(\int_0^t \left\| \nu(R_{s \wedge S_{\varepsilon,N}}) - \nu(R'_{s \wedge S_{\varepsilon,N}}) \right\|^2 ds \right) + 2T\mathbb{E}\left(\int_0^t \left\| b(R_{s \wedge S_{\varepsilon,N}}) - b(R'_{s \wedge S_{\varepsilon,N}}) \right\|^2 ds \right)
$$
$$
\leq 2(1 + T)K_{\varepsilon,N}\mathbb{E}\left(\int_0^t \left\| R_{s \wedge S_{\varepsilon,N}} - R'_{s \wedge S_{\varepsilon,N}} \right\|^2 ds \right),
$$

and then Grönwall's inequality shows

$$
\mathbb{E}\left(\| R_{t \wedge S_{\varepsilon,N}} - R'_{t \wedge S_{\varepsilon,N}} \|^2 \right) = 0.
$$

As $T > 0$ was arbitrary, this holds for all $t \geq 0$.

Next, we let $\varepsilon \to 0$. In view of the positivity (1), we have $T_\varepsilon, T'_\varepsilon \to \infty$ and conclude that

$$
R_{t \wedge S_N \wedge S'_N} = R'_{t \wedge S_N \wedge S'_N} \quad \forall t \geq 0.
$$

Together with the continuity of the paths, it follows that $S_N = S'_N$, and since this holds for all $N > 0$, we have shown that $R = R'$ up to a possible (common) time of explosion. $\square$

As weak existence together with pathwise uniqueness implies strong existence [6, Theorem IV.1.1], the results so far establish the strong wellposedness of (2-PDV) up to explosion.

Corollary 5. *The SDE* (2-PDV) *satisfies strong existence and uniqueness up to a possible explosion time.*

Turning to the main contribution of this section, we now show the absence of explosions.

Lemma 6. *If $\beta_2 < 1$ and $\lambda_1 \beta_1^2 < 2$, a solution $(R_{1,t}, R_{2,t})$ of* (2-PDV) *cannot explode in finite time. Moreover, $\sup_{t \leq T} \mathbb{E}(R_{1,t}^2 + R_{2,t}) < \infty$ for any $T \in [0, \infty)$.*

Proof. Fix $M > 0$ and define the stopping times

$$
T_M^1 := \inf\left\{ t \geq 0 : R_{1,t}^2 \geq M^2 \right\} \qquad T_M^2 := \inf\left\{ t \geq 0 : R_{2,t} \geq M^2 \right\}
$$
$$
T_M := T_M^1 \wedge T_M^2.
$$

Fix also $t \geq 0$, and note that

$$
M^2\mathbb{P}(T_M \leq t) = \mathbb{E}(M^2 \mathbf{1}_{T_M \leq t}) \leq \mathbb{E}(\max(R_{1,T_M \wedge t}^2, R_{2,T_M \wedge t})) \leq \mathbb{E}(R_{1,T_M \wedge t}^2 + R_{2,T_M \wedge t}).
$$

In the main part of the proof below, we show that

$$
\mathbb{E}(R_{1,T_M \wedge t}^2 + R_{2,T_M \wedge t}) \leq c(t) \tag{2}
$$

with $c(t) < \infty$ independent of M. It will then follow that $\lim_{M \to \infty} \mathbb{P}(T_M \leq t) = 0$, showing that R_1 and R_2 have bounded paths on any compact time interval and hence completing the proof.

To show (2), we first apply Itô's formula to obtain

$$R_{1,t\wedge T_M}^2 = R_{1,0}^2 + \int_0^{t\wedge T_M} 2\lambda_1 \sigma_s R_{1,s} dW_s + \int_0^{t\wedge T_M} (\lambda_1^2 \sigma_s^2 - 2\lambda_1 R_{1,s}^2) ds.$$

As $\sigma_s R_{1,s}$ is uniformly bounded up to the stopping time $t \wedge T_M$, it follows that

$$\mathbb{E}(R_{1,T_M\wedge t}^2) = R_{1,0}^2 + \mathbb{E}\left(\int_0^{t\wedge T_M} (\lambda_1^2 \sigma_s^2 - 2\lambda_1 R_{1,s}^2) ds\right)$$

and thus, by Fubini's theorem,

$$\mathbb{E}(R_{1,T_M\wedge t}^2) = R_{1,0}^2 + \int_0^t \mathbb{E}\left((\lambda_1^2 \sigma_s^2 - 2\lambda_1 R_{1,s}^2)\mathbf{1}_{s\leq t\wedge T_M}\right) ds. \tag{3}$$

Next, we insert the definition of σ_s^2 to get

$$\mathbb{E}(R_{1,T_M\wedge t}^2) = R_{1,0}^2 + \int_0^t \mathbb{E}\left(\left\{\lambda_1^2(\beta_0 + \beta_2\sqrt{R_{2,s}})^2\right.\right.$$
$$\left.\left. + 2\lambda_1^2\beta_1(\beta_0 + \beta_2\sqrt{R_{2,s}})R_{1,s} + (\lambda_1^2\beta_1^2 - 2\lambda_1)R_{1,s}^2\right\}\mathbf{1}_{s\leq t\wedge T_M}\right) ds.$$

We may see the expression inside the curly brackets is a quadratic polynomial $C + Bx + Ax^2$ evaluated at $x = R_{1,s}$. Our assumption that $\lambda_1\beta_1^2 < 2$ is equivalent to $A < 0$, so that the polynomial has finite maximum value

$$C - \frac{B^2}{4A} = \lambda_1^2(\beta_0 + \beta_2\sqrt{R_{2,s}})^2 + \frac{\lambda_1^4\beta_1^2(\beta_0 + \beta_2\sqrt{R_{2,s}})^2}{2\lambda_1 - \lambda_1^2\beta_1^2}$$

$$= \frac{2\lambda_1^2}{2 - \lambda_1\beta_1^2}(\beta_0 + \beta_2\sqrt{R_{2,s}})^2 \leq \frac{4\lambda_1^2\beta_0^2}{2 - \lambda_1\beta_1^2} + \frac{4\lambda_1^2\beta_2^2}{2 - \lambda_1\beta_1^2}R_{2,s}.$$

Using this as an upper bound for the expression in the curly brackets, we conclude that

$$\mathbb{E}(R_{1,T_M\wedge t}^2) \leq R_{1,0}^2 + \int_0^t \mathbb{E}\left(\left(\frac{4\lambda_1^2\beta_0^2}{2 - \lambda_1\beta_1^2} + \frac{4\lambda_1^2\beta_2^2}{2 - \lambda_1\beta_1^2}R_{2,s}\right)\mathbf{1}_{s\leq t\wedge T_M}\right) ds$$

$$\leq R_{1,0}^2 + \frac{4\lambda_1^2\beta_0^2}{2 - \lambda_1\beta_1^2}t + \frac{4\lambda_1^2\beta_2^2}{2 - \lambda_1\beta_1^2}\int_0^t \mathbb{E}(R_{2,s\wedge T_M}) ds$$

$$=: c_{1,1} + c_{1,2}t + c_{1,3}\int_0^t \mathbb{E}(R_{2,s\wedge T_M}) ds \tag{4}$$

where the constants $c_{1,1}, c_{1,2}, c_{1,3} > 0$ are independent of M and t.

Our next goal is a similar bound for R_2 instead of R_1^2. From the SDE for R_2,

$$\mathbb{E}(R_{2,T_M\wedge t})$$

$$= R_{2,0} + \lambda_2\mathbb{E}\left(\int_0^{t\wedge T_M}(\sigma_s^2 - R_{2,s}) ds\right)$$

$$= R_{2,0} + \lambda_2\int_0^t \mathbb{E}\left((\sigma_s^2 - R_{2,s})\mathbf{1}_{s\leq t\wedge T_M}\right) ds$$

$$= R_{2,0} + \lambda_2\int_0^t \mathbb{E}\left(\left\{(\beta_0 + \beta_1 R_{1,s})^2 + 2(\beta_0 + \beta_1 R_{1,s})\beta_2\sqrt{R_{2,s}} + (\beta_2^2 - 1)R_{2,s}\right\}\mathbf{1}_{s\leq t\wedge T_M}\right) ds.$$

6

Here we observe a quadratic polynomial of $\sqrt{R_{2,s}}$ inside the curly brackets, and as we have assumed that $\beta_2^2 - 1 < 0$, the polynomial has maximum value

$$(\beta_0 + \beta_1 R_{1,s})^2 + \frac{(\beta_0 + \beta_1 R_{1,s})^2 \beta_2^2}{1 - \beta_2^2} = \frac{1}{1 - \beta_2^2}(\beta_0 + \beta_1 R_{1,s})^2 \le \frac{2\beta_0^2}{1 - \beta_2^2} + \frac{2\beta_1^2}{1 - \beta_2^2} R_{1,s}^2.$$

We conclude that

$$\mathbb{E}(R_{2,T_M \wedge t}) \le R_{2,0} + \lambda_2 \int_0^t \mathbb{E}\left(\left(\frac{2\beta_0^2}{1 - \beta_2^2} + \frac{2\beta_1^2}{1 - \beta_2^2} R_{1,s}^2\right) \mathbf{1}_{s \le t \wedge T_M}\right) ds$$

$$\le R_{2,0} + \frac{2\beta_0^2 \lambda_2}{1 - \beta_2^2} t + \frac{2\beta_1^2 \lambda_2}{1 - \beta_2^2} \int_0^t \mathbb{E}\left(R_{1,s \wedge T_M}^2\right) ds$$

$$=: c_{2,1} + c_{2,2} t + c_{2,3} \int_0^t \mathbb{E}\left(R_{1,s \wedge T_M}^2\right) ds, \tag{5}$$

where again the constants do not depend on M and t.

Writing $c_i = c_{1,i} + c_{2,i}$, combining (4) and (5) yields

$$\mathbb{E}(R_{1,T_M \wedge t}^2 + R_{2,T_M \wedge t}) \le c_1 + c_2 t + c_3 \int_0^t \mathbb{E}\left(R_{1,s \wedge T_M}^2 + R_{2,s \wedge T_M}\right) ds,$$

and now Grönwall's inequality shows

$$\mathbb{E}(R_{1,T_M \wedge t}^2 + R_{2,T_M \wedge t}) \le (c_1 + c_2 t) e^{c_3 t}.$$

This establishes (2) and hence completes the proof. $\qquad\square$

2.1 Positivity of σ_t

In this section we show that $\sigma_t > 0$ under certain conditions. More precisely, we exhibit a lower bound $\sigma_t \ge Y_t > 0$ which also shows, e.g., that $1/\sigma_t$ has finite moments of all orders. This strengthens the result in [4, Section 4.1.5] where it is observed that $\sigma_t \ge 0$ since the drift of σ_t would be positive whenever σ_t reaches 0.

Theorem 7. *Consider the solution $(R_{1,t}, R_{2,t})$ of (2-PDV) up to a possible time τ of explosion. If the initial values $(R_{1,0}, R_{2,0})$ are such that $\sigma_0 = \beta_0 + \beta_1 R_{1,0} + \beta_2 \sqrt{R_{2,0}} > 0$, and if moreover $\lambda_2 < 2\lambda_1$, then $\sigma_t > 0$ for all $t < \tau$. More precisely, we have $\sigma_t \ge Y_t$, where Y_t is the stochastic exponential (9).*

Proof. By Itô's formula, σ_t satisfies the SDE

$$d\sigma_t = \left(-\beta_1 \lambda_1 R_{1,t} + \frac{\lambda_2 \beta_2}{2} \frac{\sigma_t^2 - R_{2,t}}{\sqrt{R_{2,t}}}\right) dt + \beta_1 \lambda_1 \sigma_t dW_t. \tag{6}$$

Using the assumption that $\lambda_2 < 2\lambda_1$, we can bound the drift of σ_t from below:

$$-\beta_1 \lambda_1 R_{1,t} + \frac{\lambda_2 \beta_2}{2} \frac{\sigma_t^2 - R_{2,t}}{\sqrt{R_{2,t}}} = -\lambda_1(\sigma_t - \beta_0 - \beta_2 \sqrt{R_{2,t}}) - \frac{\lambda_2 \beta_2}{2} \sqrt{R_{2,t}} + \frac{\lambda_2 \beta_2}{2} \frac{\sigma_t^2}{\sqrt{R_{2,t}}}$$

$$\ge -\lambda_1 \sigma_t + \beta_0 \lambda_1 + \beta_2 \left(\lambda_1 - \frac{\lambda_2}{2}\right) \sqrt{R_{2,t}} + \frac{\lambda_2 \beta_2}{2} \frac{\sigma_t^2}{\sqrt{R_{2,t}}}$$

$$\ge -\lambda_1 \sigma_t \tag{7}$$

because all the other terms are nonnegative. Inspired by (7), we define a process Y via

$$dY_t = -\lambda_1 Y_t dt + \beta_1 \lambda_1 Y_t dW_t , \quad Y_0 = \sigma_0. \tag{8}$$

Note that Y is simply the stochastic exponential

$$Y_t = \sigma_0 \exp\left(\beta_1 \lambda_1 W_t - \lambda_1 t - \frac{1}{2}\beta_1^2 \lambda_1^2 t\right) \tag{9}$$

and, in particular, $Y_t > 0$ for all $t \geq 0$.

The SDEs (6) and (8) have the same initial condition σ_0, the same volatility function $v(x) = \beta_1 \lambda_1 x$ and their drift functions are ordered according to (7). Moreover, both drift and volatility functions are continuous on the relevant domains, the drift function of Y is Lipschitz, and

$$\int_0^\varepsilon v(x)^{-2} dx = \infty$$

for every $\varepsilon > 0$. In view of these conditions, the comparison result for SDEs[1] [7, Theorem 5.2.18] yields that $\sigma_t \geq Y_t$ for all $t < \tau$. □

3   Analysis of the 4-Factor Model (4-PDV)

The general wellposedness of (4-PDV) is shown using the same arguments as in the 2-factor model; we therefore omit the proof.

Proposition 8. *The SDE* (4-PDV) *satisfies strong existence and uniqueness up to a possible explosion time.*

The next result contains our main contribution.

Lemma 9. *Suppose that*

$$\beta_2^2 \theta_2 < 1, \qquad \beta_2^2(1-\theta_2) < 1, \qquad \lambda_{1,0}\beta_1^2(1-\theta_1)^2 < 2, \qquad \lambda_{1,1}\beta_1^2\theta_1^2 < 2.$$

Then a solution $(R_{1,j,t}, R_{2,j,t})_{j\in\{0,1\}}$ *of* (4-PDV) *cannot explode in finite time. Moreover,* $\sup_{t\leq T}\mathbb{E}(R_{1,0,t}^2 + R_{1,1,t}^2 + R_{2,0,t}^2 + R_{2,1,t}^2) < \infty$ *for any* $T \in [0, \infty)$.

Proof. We follow the guidance provided by the 2-factor model. Fix $M > 0$ and define

$$T_M^{1,0} := \inf\left\{t \geq 0 : R_{1,0,t}^2 \geq M^2\right\} \qquad T_M^{1,1} := \inf\left\{t \geq 0 : R_{1,1,t}^2 \geq M^2\right\}$$
$$T_M^{2,0} := \inf\{t \geq 0 : R_{2,0,t}^2 \geq M^2\} \qquad T_M^{2,1} := \inf\{t \geq 0 : R_{2,1,t}^2 \geq M^2\}$$
$$T_M := T_M^{1,0} \wedge T_M^{1,1} \wedge T_M^{2,0} \wedge T_M^{2,1}.$$

---

[1]To be precise, the cited theorem is stated for SDEs where the drift and volatility functions depend only on time and the solution process. Here, they are random as they depend on $(R_{1,t}, R_{2,t})$. The proof holds without change.

Fix also $t > 0$, then

$$
\begin{aligned}
M^2 \mathbb{P}(T_M \le t) = \mathbb{E}(M^2 \mathbf{1}_{T_M \le t}) \\
\le \mathbb{E}\left(\max\left(R^2_{1,0,t \wedge T_M}, R^2_{1,1,t \wedge T_M}, R_{2,0,t \wedge T_M}, R_{2,1,t \wedge T_M}\right)\right) \\
\le \mathbb{E}\left(R^2_{1,0,t \wedge T_M} + R^2_{1,1,t \wedge T_M} + R_{2,0,t \wedge T_M} + R_{2,1,t \wedge T_M}\right) = \mathbb{E}(U_{t \wedge T_M})
\end{aligned}
$$

where

$$
U_t := R^2_{1,0,t} + R^2_{1,1,t} + R_{2,0,t} + R_{2,1,t} \,.
$$

We shall prove that

$$
\mathbb{E}(U_{t \wedge T_M}) \le C(t) \tag{10}
$$

for a continuous $C(t)$ independent of M, and that will imply the claim.

To prove (10), we note as in (3) that

$$
\mathbb{E}\left(R^2_{1,j,t \wedge T_M}\right) = R^2_{1,j,0} + \int_0^t \mathbb{E}\left(\{\lambda^2_{1,j}\sigma^2_s - 2\lambda_{1,j}R^2_{1,j,s}\}\mathbf{1}_{s \le t \wedge T_M}\right) ds. \tag{11}
$$

We first focus on $j = 0$. Inserting the definition of σ^2_s and then the one of $R^2_{1,s}$, we have

$$
\begin{aligned}
\lambda^2_{1,0}\sigma^2_s - & 2\lambda_{1,0}R^2_{1,0,s} \\
&= \lambda^2_{1,0}(\beta_0 + \beta_2\sqrt{R_{2,s}})^2 + 2\lambda^2_{1,0}\beta_1(\beta_0 + \beta_2\sqrt{R_{2,s}})R_{1,s} + \lambda^2_{1,0}\beta^2_1 R^2_{1,s} - 2\lambda_{1,0}R^2_{1,0,s} \\
&= C + BR_{1,0,s} + AR^2_{1,0,s}
\end{aligned}
$$

where

$$
\begin{aligned}
C &= \lambda^2_{1,0}(\beta_0 + \beta_2\sqrt{R_{2,s}})^2 + 2\lambda^2_{1,0}\beta_1\theta_1(\beta_0 + \beta_2\sqrt{R_{2,s}})R_{1,1,s} + \lambda^2_{1,0}\beta^2_1\theta^2_1 R^2_{1,1,s} \\
B &= 2\lambda^2_{1,0}\beta_1(1 - \theta_1)(\beta_0 + \beta_2\sqrt{R_{2,s}}) + 2\lambda^2_{1,0}\beta^2_1\theta_1(1 - \theta_1)R_{1,1,s} \\
A &= \lambda^2_{1,0}\beta^2_1(1 - \theta_1)^2 - 2\lambda_{1,0} \,.
\end{aligned}
$$

Our assumption that $\lambda_{1,0}\beta^2_1(1 - \theta_1)^2 < 2$ means that $A < 0$, hence the quadratic polynomial $C + Bx + Ax^2$ has a maximum. By an elementary if tedious calculation, that value is

$$
C - \frac{B^2}{4A} = \frac{2\lambda^2_{1,0}}{2 - \lambda_{1,0}\beta^2_1(1 - \theta_1)^2}\left(\beta_0 + \beta_2\sqrt{R_{2,s}} + \beta_1\theta_1 R_{1,1,s}\right)^2.
$$

Applying the elementary bound $(a + b)^2 \le 2a^2 + 2b^2$ twice, we conclude

$$
\lambda^2_{1,0}\sigma^2_s - 2\lambda_{1,0}R^2_{1,0,s} \le \frac{4\lambda^2_{1,0}}{2 - \lambda_{1,0}\beta^2_1(1 - \theta_1)^2}\left(2\beta^2_0 + 2\beta^2_2 R_{2,s} + \beta^2_1\theta^2_1 R^2_{1,1,s}\right).
$$

Using this inequality in (11) yields

$$
\begin{aligned}
\mathbb{E}\left(R^2_{1,0,t \wedge T_M}\right) &\le R^2_{1,0,0} + \int_0^t \mathbb{E}\left(\frac{4\lambda^2_{1,0}}{2 - \lambda_{1,0}\beta^2_1(1 - \theta_1)^2}\left(2\beta^2_0 + 2\beta^2_2 R_{2,s \wedge T_M} + \beta^2_1\theta^2_1 R^2_{1,1,s \wedge T_M}\right)\right) ds \\
&=: c_{1,0,0}(t) + c_{1,0,1}\int_0^t \mathbb{E}(R_{2,s \wedge T_M})ds + c_{1,0,2}\int_0^t \mathbb{E}(R^2_{1,1,s \wedge T_M})ds
\end{aligned}
$$

9

where $c_{1,0,0}(t)$ is affine in t.

Symmetrically, we have for $j = 1$ that

$$\lambda_{1,1}^2 \sigma_s^2 - 2\lambda_{1,1} R_{1,1,s}^2 \leq \frac{4\lambda_{1,1}^2}{2 - \lambda_{1,1}\beta_1^2\theta_1^2} \left(2\beta_0^2 + 2\beta_2^2 R_{2,s} + \beta_1^2(1 - \theta_1)^2 R_{1,0,s}^2\right)$$

and then

$$\mathbb{E}\left(R_{1,1,t\wedge T_M}^2\right) \leq R_{1,1,0}^2 + \int_0^t \mathbb{E}\left(\frac{4\lambda_{1,1}^2}{2 - \lambda_{1,1}\beta_1^2\theta_1^2}\left(2\beta_0^2 + 2\beta_2^2 R_{2,s\wedge T_M} + \beta_1^2(1 - \theta_1)^2 R_{1,0,s\wedge T_M}^2\right)\right) ds$$

$$:= c_{1,1,0}(t) + c_{1,1,1}\int_0^t \mathbb{E}(R_{2,s\wedge T_M}) ds + c_{1,1,2}\int_0^t \mathbb{E}(R_{1,0,s\wedge T_M}^2) ds.$$

Combining the results for $j = 0$ and $j = 1$ yields

$$\mathbb{E}\left(R_{1,0,t\wedge T_M}^2 + R_{1,1,t\wedge T_M}^2\right)$$

$$\leq c_{1,0}(t) + c_{1,1}\int_0^t \mathbb{E}(R_{2,s\wedge T_M}) ds + c_{1,2}\int_0^t \mathbb{E}\left(R_{1,0,s\wedge T_M}^2 + R_{1,1,s\wedge T_M}^2\right) ds \quad (12)$$

where the constants have the obvious definitions.

Next, we derive a similar bound for $R_{2,j}$ instead of $R_{1,j}^2$. From the SDE for $R_{2,j}$,

$$\mathbb{E}\left(R_{2,j,t\wedge T_M}\right) = R_{2,j,0} + \lambda_{2,j}\int_0^t \mathbb{E}\left((\sigma_s^2 - R_{2,j,s})\mathbf{1}_{s\leq t\wedge T_M}\right) ds. \quad (13)$$

Focusing again on $j = 0$ first, we insert the definitions of σ_s^2 and $R_{2,s}$ and estimate

$$\sigma_s^2 - R_{2,0,s}$$
$$= (\beta_0 + \beta_1 R_{1,s})^2 + 2\beta_2(\beta_0 + \beta_1 R_{1,s})\sqrt{R_{2,s}} + \beta_2^2\theta_2 R_{2,1,s} + (\beta_2^2(1 - \theta_2) - 1)R_{2,0,s}$$
$$\leq (\beta_0 + |\beta_1||R_{1,s}|)^2 + 2\beta_2(\beta_0 + |\beta_1||R_{1,s}|)\sqrt{R_{2,s}} + \beta_2^2\theta_2 R_{2,1,s} + (\beta_2^2(1 - \theta_2) - 1)R_{2,0,s}$$
$$\leq (\beta_0 + |\beta_1||R_{1,s}|)^2 + 2\beta_2(\beta_0 + |\beta_1||R_{1,s}|)(\sqrt{\theta_2}\sqrt{R_{2,1,s}} + \sqrt{1 - \theta_2}\sqrt{R_{2,0,s}})$$
$$\quad + \beta_2^2\theta_2 R_{2,1,s} + (\beta_2^2(1 - \theta_2) - 1)R_{2,0,s}$$
$$= C + B\sqrt{R_{2,0,s}} + AR_{2,0,s}$$

where

$$A = \beta_2^2(1 - \theta_2) - 1$$
$$B = 2\beta_2\sqrt{1 - \theta_2}(\beta_0 + |\beta_1||R_{1,s}|)$$
$$C = (\beta_0 + |\beta_1||R_{1,s}|)^2 + \beta_2^2\theta_2 R_{2,1,s} + 2\beta_2\sqrt{\theta_2}(\beta_0 + |\beta_1||R_{1,s}|)\sqrt{R_{2,1,s}}.$$

The polynomial $C + Bx + Ax^2$ has a maximum as $A < 0$ due to $\beta_2^2(1 - \theta_2) < 1$, and the maximum value is

$$C - \frac{B^2}{4A} = \frac{1}{1 - \beta_2^2(1 - \theta_2)}(\beta_0 + |\beta_1||R_{1,s}|)^2 + \beta_2^2\theta_2 R_{2,1,s} + 2\beta_2\sqrt{\theta_2}(\beta_0 + |\beta_1||R_{1,s}|)\sqrt{R_{2,1,s}}.$$

10

It follows that

$$\sigma_s^2 - R_{2,0,s} \leq \frac{1}{1 - \beta_2^2(1 - \theta_2)}(\beta_0 + |\beta_1||R_{1,s}|)^2 + \beta_2^2\theta_2 R_{2,1,s} + 2\beta_2\sqrt{\theta_2}(\beta_0 + |\beta_1||R_{1,s}|)\sqrt{R_{2,1,s}}$$

$$\leq \frac{1}{1 - \beta_2^2(1 - \theta_2)}(\beta_0 + |\beta_1||R_{1,s}|)^2 + \beta_2^2\theta_2 R_{2,1,s} + \beta_2^2\theta_2 R_{2,1,s} + (\beta_0 + |\beta_1||R_{1,s}|)^2$$

$$= \frac{2 - \beta_2^2(1 - \theta_2)}{1 - \beta_2^2(1 - \theta_2)}(\beta_0 + |\beta_1||R_{1,s}|)^2 + 2\beta_2^2\theta_2 R_{2,1,s}$$

$$\leq \frac{2 - \beta_2^2(1 - \theta_2)}{1 - \beta_2^2(1 - \theta_2)}(2\beta_0^2 + 2\beta_1^2 R_{1,s}^2) + 2\beta_2^2\theta_2 R_{2,1,s}.$$

Applying this in (13), we conclude that

$$\mathbb{E}\left(R_{2,0,t\wedge T_M}\right) \leq R_{2,0,0} + \lambda_{2,0}\int_0^t \mathbb{E}\left(\frac{2 - \beta_2^2(1 - \theta_2)}{1 - \beta_2^2(1 - \theta_2)}(2\beta_0^2 + 2\beta_1^2 R_{1,s\wedge T_M}^2) + 2\beta_2^2\theta_2 R_{2,1,s\wedge T_M}\right)ds$$

$$=: c_{2,0,0}(t) + c_{2,0,1}\int_0^t \mathbb{E}\left(R_{1,s\wedge T_M}^2\right)ds + c_{2,0,2}\int_0^t \mathbb{E}\left(R_{2,1,s\wedge T_M}\right)ds.$$

Symmetrically, we obtain for $j = 1$ that

$$\sigma_s^2 - R_{2,1,s} \leq \frac{2 - \beta_2^2\theta_2}{1 - \beta_2^2\theta_2}(2\beta_0^2 + 2\beta_1^2 R_{1,s}^2) + 2\beta_2^2(1 - \theta_2)R_{2,0,s}$$

and then

$$\mathbb{E}\left(R_{2,1,t\wedge T_M}\right) \leq R_{2,1,0} + \lambda_{2,0}\int_0^t \mathbb{E}\left(\frac{2 - \beta_2^2\theta_2}{1 - \beta_2^2\theta_2}(2\beta_0^2 + 2\beta_1^2 R_{1,s\wedge T_M}^2) + 2\beta_2^2(1 - \theta_2)R_{2,0,s\wedge T_M}\right)ds$$

$$:= c_{2,1,0}(t) + c_{2,1,1}\int_0^t \mathbb{E}\left(R_{1,s\wedge T_M}^2\right)ds + c_{2,1,2}\int_0^t \mathbb{E}\left(R_{2,0,s\wedge T_M}\right)ds.$$

Adding the two inequalities, we deduce

$$\mathbb{E}\left(R_{2,0,t\wedge T_M} + R_{2,1,t\wedge T_M}\right)$$

$$\leq c_{2,0}(t) + c_{2,1}\int_0^t \mathbb{E}(R_{1,s\wedge T_M}^2)ds + c_{2,2}\int_0^t \mathbb{E}\left(R_{2,0,s\wedge T_M} + R_{2,1,s\wedge T_M}\right)ds$$

$$\leq c_{2,0}(t) + c_{2,1}\int_0^t \mathbb{E}(R_{1,0,s\wedge T_M}^2 + R_{1,1,s\wedge T_M}^2)ds + c_{2,2}\int_0^t \mathbb{E}\left(R_{2,0,s\wedge T_M} + R_{2,1,s\wedge T_M}\right)ds$$

$$\tag{14}$$

where we used the elementary convexity inequality

$$R_{1,s}^2 = ((1 - \theta_1)R_{1,0,s} + \theta_1 R_{1,1,s})^2 \leq (1 - \theta_1)R_{1,0,s}^2 + \theta_1 R_{1,1,s}^2 \leq R_{1,0,s}^2 + R_{1,1,s}^2.$$

We can now add (12) and (14) to obtain

$$\mathbb{E}(U_{t\wedge T_M}) \leq c_0(t) + c_1\int_0^t \mathbb{E}(U_{s\wedge T_M})ds$$

where $c_0(t)$ is affine and nondecreasing in t and $c_0(t), c_1$ do not depend on M. Grönwall's inequality then yields

$$\mathbb{E}(U_{t\wedge T_M}) \leq c_0(t)e^{c_1 t}$$

which is the desired bound (10). $\qquad\square$

11

3.1 Failure of Positivity of σ_t in (4-PDV)

In [4] it is reported that for realistic parameter values, the volatility σ_t remained positive in simulations of the 4-factor model. Nevertheless, existence of a reasonable sufficient condition for $\sigma_t > 0$ (or even just $\sigma_t \geq 0$) in (4-PDV) remains open. Below, we explain that a direct generalization of Theorem 7 fails. Indeed, in the 4-factor model, σ_t follows the SDE

$$d\sigma_t = \left(-\beta_1\bar{\lambda}_1\bar{R}_{1,t} + \frac{\bar{\lambda}_2\beta_2}{2}\frac{\sigma_t^2 - \bar{R}_{2,t}}{\sqrt{R_{2,t}}} \right) dt + \beta_1\bar{\lambda}_1\sigma_t dW_t \tag{15}$$

where

$$\bar{\lambda}_i := (1-\theta_i)\lambda_{i,0} + \theta_i\lambda_{i,1}, \qquad \bar{R}_{i,t} := \frac{(1-\theta_i)\lambda_{i,0}R_{i,0,t} + \theta_i\lambda_{i,1}R_{i,1,t}}{\bar{\lambda}_i}$$

as seen in [4]. The analogue of the condition in Theorem 7 is $\bar{\lambda}_2 < 2\bar{\lambda}_1$.

Proposition 10. *Under the conditions of Theorem 2 for wellposedness, and even if $\bar{\lambda}_2 < 2\bar{\lambda}_1$, it may happen that $\sigma_0 > 0$ but $\mathbb{P}(\sigma_t < 0) > 0$ for some $t > 0$.*

Proof. We choose initial conditions $R_{1,0,0} < 0$ and $R_{1,1,0} > 0$, and coefficients $\theta_1, \lambda_{1,0}, \lambda_{1,1}$, such that $R_{1,0} > 0$ and $\bar{R}_{1,0} < 0$. Next, choose $\beta_1 = -1$ (say), and $\beta_2 > 0$ small enough such that $\beta_1 R_{1,0} + \beta_2\sqrt{R_{2,0}} = 0$. Consider for the moment $\beta_0 := 0$, then the preceding identity means that $\sigma_0 = 0$. Inspecting (15), we see that at $t = 0$, the volatility vanishes while the drift rate is

$$-\beta_1\bar{\lambda}_1\bar{R}_{1,t} - \frac{\bar{\lambda}_2\beta_2}{2}\frac{\bar{R}_{2,t}}{\sqrt{R_{2,t}}} < 0.$$

By continuity of the paths, it follows that $\mathbb{P}(\sigma_t < 0) > 0$ for all $t > 0$ sufficiently small.

Next, we modify the above by choosing β_0 strictly positive, so that $\sigma_0 = \beta_0 > 0$. We may see β_0 as a parameter of the SDE determining σ_t. If the solution is continuous with respect to β_0, it follows that $\mathbb{P}(\sigma_t < 0) > 0$ for $\beta_0 > 0$ and t sufficiently small. Continuity is a standard result for SDEs with Lipschitz coefficients (e.g., [8, Section 4.5]). To see that the Lipschitz result is sufficient, note that for the present purpose of showing that $\mathbb{P}(\sigma_t < 0) > 0$ for some small $t > 0$, we may truncate the non-Lipschitz coefficients in (4-PDV); that is, we replace $\sqrt{R_{2,t}}$ by $\sqrt{R_{2,t} \vee \delta}$ and σ_t^2 by $\sigma_t^2 \wedge \delta^{-1}$ for a small constant $\delta > 0$. $\qquad\square$

References

[1] R. Chicheportiche and J.-P. Bouchaud. The fine-structure of volatility feedback I: multi-scale self-reflexivity. *Phys. A*, 410:174–195, 2014.

[2] P. Foschi and A. Pascucci. Path dependent volatility. *Decis. Econ. Finance*, 31(1):13–32, 2008.

[3] J. Guyon. Path-dependent volatility. *Risk*, October 2014.

[4] J. Guyon and J. Lekeufack. Volatility is (mostly) path-dependent. *Preprint SSRN*, 2022. Available at https://ssrn.com/abstract=4174589.

[5] D. G. Hobson and L. C. G. Rogers. Complete models with stochastic volatility. *Math. Finance*, 8(1):27–48, 1998.

[6] N. Ikeda and S. Watanabe. *Stochastic differential equations and diffusion processes*, volume 24 of *North-Holland Mathematical Library*. North-Holland, 2nd edition, 1989.

[7] I. Karatzas and S. E. Shreve. *Brownian Motion and Stochastic Calculus*. Springer, New York, 2nd edition, 1991.

[8] H. Kunita. *Stochastic flows and stochastic differential equations*, volume 24 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1990.

[9] G. Zumbach. Volatility conditional on price trends. *Quant. Finance*, 10(4):431–442, 2010.