DOES THE TERM-STRUCTURE OF EQUITY AT-THE-MONEY SKEW REALLY FOLLOW A POWER LAW?

MEHDI EL AMRANI

Bloomberg L.P., Quantitative Research

JULIEN GUYON

Bloomberg L.P., Quantitative Research
Ecole des Ponts ParisTech, CERMICS

Abstract. Using two years of S&P 500, Eurostoxx 50, and DAX data, we empirically investigate the term-structure of the at-the-money-forward (ATM) skew of equity indexes. While a power law (2 parameters) captures the term-structure well away from short maturities, the power law fit deteriorates considerably when short maturities are included. By contrast, 3-parameter shapes that look like power laws but do not blow up at vanishing maturity, such as time-shifted or capped power laws, are shown to fit well regardless of whether short maturities are included or not. Our study suggests that the term-structure of equity ATM skew has a power-law shape for maturities above 1 month but has a different behavior, and in particular may not blow up, for shorter maturities. The 3-parameter shapes are derived from non-Markovian variance curve models using the Bergomi-Guyon expansion. A simple 4-parameter term-structure similarly derived from the (Markovian) two-factor Bergomi model is also considered and provides even better fits. The extrapolated zero-maturity skew, far from being infinite, is distributed around a typical value of 1.5 (in absolute value).

Keywords. At-the-money skew, equity skew, variance curve models, calibration, power law, rough volatility, Bergomi model

1. Introduction

Many articles, including for instance [6, 9, 11, 10, 2, 8], convey a message that the term-structure of S&P 500 (SPX) at-the-money-forward (ATM) skew follows a power-law decay $\mathrm{PL}^*_{c,\alpha}(T) := cT^{-\alpha}$, with $\alpha \in (0.3, 0.5)$, and in particular blows up when the maturity T vanishes. Is this really true? Do other simple parametric shapes, not blowing up for vanishing maturities, actually fit better this term-structure, consistently over time and across equity indexes? To what volatility models do these simple parametric shapes correspond? Those are the questions we investigate in this article. Note that the ATM skew of equity indexes is negative. Since it is more convenient to work with positive quantities, **throughout this article, by ATM skew we mean the negative (or absolute value) of the actual, negative skew**.

In [12, Section 4.4] it was shown that the term-structure of SPX ATM skew is much better captured by the 4-parameter shape $\mathrm{B}_{k_1,k_2,c_1,c_2}(T)$ given by the Bergomi-Guyon expansion [4] of the two-factor Bergomi (2fB) model [3] than by a power law, at least on the day when the calibration was performed (October 8, 2019). (The Bergomi-Guyon expansion is a general expansion of the smile of stochastic volatility models at second order in the volatility-of-volatility.) To be more precise, for that calibration date, Figure 1.1 shows that *when we ignore the shortest monthly market maturity of SPX options*, (a) a $T^{-0.47}$ power law does indeed fit the market data well, and (b) the 2fB shape, $\mathrm{B}_{k_1,k_2,c_1,c_2}(T)$, which has two more parameters (4: k_1, k_2, c_1, c_2, versus 2: c, α), fits only slightly better, mostly for long maturities. On that date, both fits are quite close to each other, and they both greatly overestimate the shortest SPX ATM skew, by around 50%.

However, Figure 1.1 also shows that when we include the first monthly market maturity of SPX options, the new best power law ($T^{-0.36}$) fails to fit the term-structure of SPX ATM skew, being too low for short

E-mail addresses: melamranizi1@bloomberg.net, julien.guyon@enpc.fr.
Date: First version: February 17, 2022. This version: March 27, 2023.

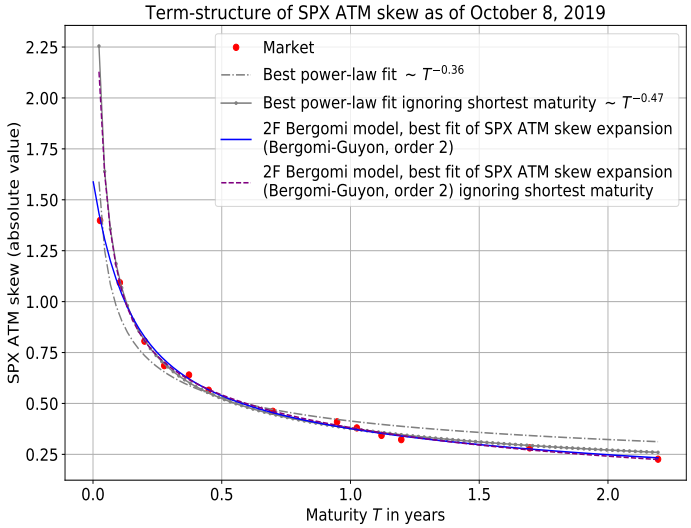

FIGURE 1.1. SPX ATM skew as of October 8, 2019. Comparison of power-law fits and fits using the second-order Bergomi-Guyon expansion of the SPX ATM skew in the 2fB model

maturities $T \in [0.05, 0.5]$ and too large for long maturities $T \geq 0.9$. By contrast, thanks to its greater flexibility, the 4-parameter shape $\mathrm{B}_{k_1,k_2,c_1,c_2}(T)$ is able to fit the new market data point without damaging the whole calibration. Of course one calibration on one given date is not enough to draw conclusions; the main contribution of this article is precisely to perform a systematic empirical study covering two years of calibration dates and several major equity indexes. Note that [11] also reports a power-law fit that is way too large at short maturities, where the data points do not indicate any blow up.

The power-law decay corresponds to the short maturity asymptotics of rough volatility models [1, 7, 10, 2], such as the rough Bergomi model. The fact that the term-structure of SPX ATM skew exhibits a power-law-like decay for a large range of maturities (as we argue, at least away from very short maturities) has been seen as a sign that volatility is rough, and it seems that many authors now take for granted that market ATM skews of equity indexes blow up for very short maturities, like in rough volatility models. In fact, our empirical study, which analyzes two years of S&P 500, Eurostoxx 50, and DAX data, shows that while a power law (2 parameters) captures the term-structure well away from short maturities, say, for maturities from 1 month to 3 years, the power law fit deteriorates considerably when short maturities are included (see Figure 4.10). In particular, it deteriorates quickly as the first monthly options maturity T_1 gets closer (see Figure 4.3). By contrast, 3-parameter shapes that globally look like power laws but do not blow up at vanishing maturity (think of a power law with a small time shift to the left, or a capped power law) are shown to fit well whether short maturities are included or not. In particular, their goodness of fit does not deteriorate as T_1 gets closer (see Figures 4.3 and 4.10). This strongly suggests that, while the term-structure of equity ATM skew indeed has a power-law shape for a large range of maturities from, say, 1 month to 3 years, it actually may not blow up for very short maturities. In fact, plotting average log-ATM skews against log-maturity suggests that ATM skews follow two distinct regimes, one for maturities below 3 weeks, and one for maturities above 3 weeks (see Figure 4.12). For Eurostoxx and DAX, the average ATM skew actually seems to follow a capped power law. We show that the extrapolated zero-maturity skew, far from being infinite, is distributed around 1.5 (in absolute value, Figure 4.14).

The 3-parameter shapes are derived from simple non-Markovian variance curve models using the Bergomi-Guyon expansion. The simple 4-parameter term-structure $\mathrm{B}_{k_1,k_2,c_1,c_2}(T)$, similarly derived from the 2fB model, is also considered and, having one more parameter, provides even better fits. The great benefit of the 2fB model is that it is Markovian, hence much easier and faster to simulate than rough volatility or other non-Markovian models. It also very naturally captures the short and long time scale properties of volatility implied by the term-structure of ATM skew. From Figure 4.7, the typical fast and slow time scales of the 2fB model are $365/k_1 \approx 10$ days and $365/k_2 \approx 200$ days.

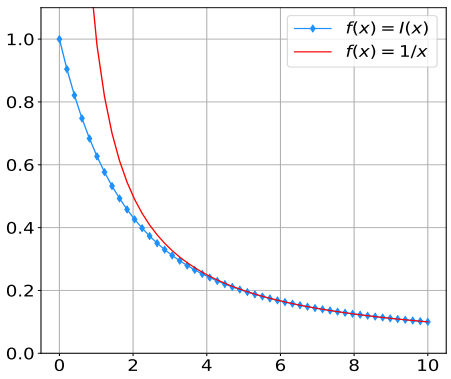

FIGURE 2.1. Graph of the functions I and $x \mapsto 1/x$

Our conclusions are in line with the recent work [15] who also finds that "the fractional kernel as used in the rough Bergomi and rough Heston models is not flexible enough to separate the short and long time scale properties of volatility that is implied by quoted SPX options;" that "the fractional kernel [...] lacks flexibility in decoupling the short and long lag volatility autocorrelations;" and that "the volatility autocorrelation structure is better captured by a classical but two-factor volatility model."

The remainder of this article is structured as follows. Section 2 presents the parametrizations of the term-structure of ATM skew that we consider in our study. In Section 3 we describe the data and our calibration methodology. The results are presented and discussed in Section 4. Finally, Section 5 concludes.

2. Parametrizations of the ATM skew

General variance curve models are described by the following dynamics

$$\begin{aligned} \frac{dS_t}{S_t} &= (r_t - q_t)\,dt + \sqrt{\xi_t^t}\,dW_t^1 \\ d\xi_t^u &= \lambda(t,u,\xi_t)\cdot dW_t, \qquad u>0, \quad t\in[0,u) \end{aligned} \tag{2.1}$$

where $W=(W^1,\dots,W^d)$ is a d-dimensional standard Brownian motion, r_t is the instantaneous interest rate, q_t is the continuous dividend yield and $\xi_t^\cdot$ denotes the instantaneous forward variance curve. The Bergomi-Guyon expansion of the ATM skew [4] (at second order in volatility-of-volatility ω) only depends on the first component $\Lambda := \lambda_1$ of the vector-valued function λ, the instantaneous spot-variance covariance function, which encapsulates all the spot-vol covariance information:

$$\mathcal{S}(T) = \sigma_{\text{VS}}\left(\frac{C^{x\xi}}{2v^2} + \frac{4C^{\mu}v - 3(C^{x\xi})^2}{8v^3}\right) + O(\omega^3), \tag{2.2}$$

where $v := \int_0^T \xi_0^u\,du$ denotes the integrated variance over $[0,T]$ seen from time 0, $\sigma_{\text{VS}} := \sqrt{v/T}$ is the corresponding variance swap rate, and, assuming that $\Lambda(t,u,\xi_t^\cdot)$ depends on $\xi_t^\cdot$ only through ξ_t^u,

$$\begin{aligned} C^{x\xi} &:= \int_0^T du \int_0^u dt\,\sqrt{\xi_0^t}\,\Lambda(t,u,\xi_0^u), \\ C^{\mu} &:= \int_0^T ds \int_s^T du\,\sqrt{\xi_0^s}\,\Lambda(s,u,\xi_0^u)\left(\frac{1}{2\sqrt{\xi_0^u}}\int_u^T dt\,\Lambda(u,t,\xi_0^t) + \int_s^u dr\,\sqrt{\xi_0^r}\,\partial_3\Lambda(r,u,\xi_0^u)\right). \end{aligned}$$

For example, for the 2fB model, Λ is of the form

$$\Lambda_{\text{2fB}}(t,u,y) = \left(c_1 e^{-k_1(u-t)} + c_2 e^{-k_2(u-t)}\right) y^u, \qquad c_1,c_2\in\mathbb{R}, \quad k_1,k_2>0 \tag{2.3}$$

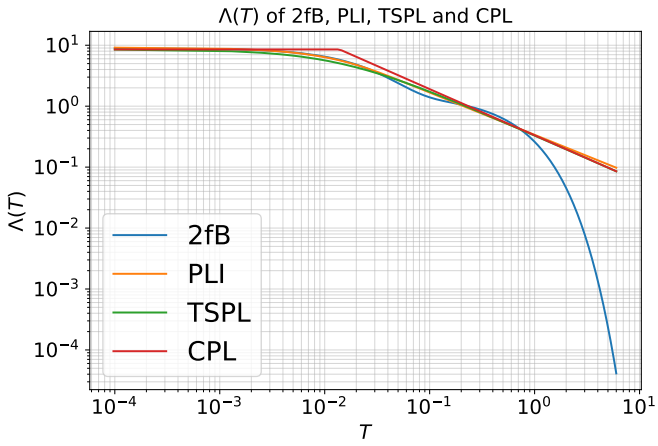

FIGURE 2.2. $\Lambda(T)$ of 2fB, PLI, TSPL, and CPL. 2fB parameters: $c_1 = 7.5, c_2 = 1.5, k_1 = 40, k_2 = 1.75, y^u = 1$

where y denotes the curve $(y^s)_{t\le s\le u}$. In addition to the 2fB model, we will also consider variance curve models in which Λ takes one of the following forms:

$$\text{rough Bergomi:} \quad \Lambda_{\mathrm{rB}}(t,u,y) := c(u-t)^{-\alpha}y^u, \qquad c\in\mathbb{R}, \quad \alpha\in(0,1/2) \tag{2.4}$$

$$\text{Capped Power Law:} \quad \Lambda_{\mathrm{CPL}}(t,u,y) := c\max(u-t,\delta)^{-\alpha}y^u, \qquad c\in\mathbb{R}, \quad \alpha,\delta>0 \tag{2.5}$$

$$\text{Time-Shifted Power Law:} \quad \Lambda_{\mathrm{TSPL}}(t,u,y) := c(u-t+\delta)^{-\alpha}y^u, \qquad c\in\mathbb{R}, \quad \alpha,\delta>0 \tag{2.6}$$

$$\text{Power Law of Function } I\text{:} \quad \Lambda_{\mathrm{PLI}}(t,u,y) := cI(k(u-t))^{\alpha}y^u, \qquad c\in\mathbb{R}, \quad \alpha,k>0 \tag{2.7}$$

where the function I is defined by

$$I(x) := \frac{1-e^{-x}}{x}, \quad x>0, \qquad I(0) := 1. \tag{2.8}$$

The graph of the function I is given in Figure 2.1. These instantaneous spot-variance covariance functions correspond for instance to the following Gaussian variance curve models (here we write the dynamics for one-factor models, i.e., $d = 2$, but this is only for simplicity)

$$\frac{d\xi^u_t}{\xi^u_t} = K(u-t)(\rho\, dW^1_t + \sqrt{1-\rho^2}\, dW^2_t) \tag{2.9}$$

with the convolution kernel K being one of the following:

$$K_{\mathrm{rB}}(\tau) = \omega\tau^{-\alpha}, \quad K_{\mathrm{CPL}}(\tau) := \omega\max(\tau,\delta)^{-\alpha}, \quad K_{\mathrm{TSPL}}(\tau) := \omega(\tau+\delta)^{-\alpha}, \quad K_{\mathrm{PLI}}(\tau) := \omega I(k\tau)^{\alpha}. \tag{2.10}$$

K_{rB} is the rough Bergomi (rB) kernel [2]; K_{CPL} is the capped power-law (CPL) kernel; K_{TSPL} is the time-shifted power-law (TSPL) kernel; the last kernel K_{PLI} compounds the function I and the power law (PL) in this order. The last two kernels have been suggested in [13, Section 4.3] precisely in order to reproduce the power-law like decay (away from very short maturities) of the term-structure of ATM skew, including for very large maturities, without generating a blow-up for vanishing maturities; K_{CPL} is similar. The CPL, TSPL, and PLI Λs have only one extra parameter (3 in total) compared to the rough Bergomi Λ (2 parameters). Λ_{2fB} has one more parameter (4 in total) than the 3-parameter Λs. Since the 3-parameter Λs and Λ_{2fB} approximate each other well on the 0-1Y maturity range (see Figure 2.2), we expect all these models to fit the term-structure of ATM skew equally well, at least for maturities not too large.[1] The main advantage of working with the 2fB model is that it is a Markovian model which is much easier and faster to simulate than the non-Markovian models (2.9)-(2.10). We do not consider the one-factor Bergomi model ($c_2 = 0$, 2 parameters k_1, c_1) as it is well known that it very poorly fits the term-structure of equity ATM skew.

[1]Note that, as proved in [4], the exponential tail of the 2fB Λ translates into an asymptotic T^{-1} power-law decay—not an exponential one—of long-term skew (for $T \gg 1/k_2$), while power-law tails with exponent α imply a long-term skew decaying as $T^{-(\alpha\wedge 1)}$.

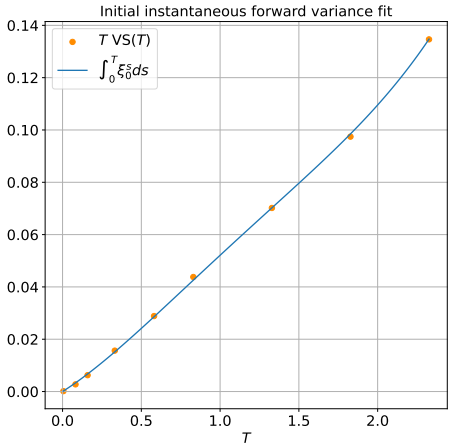

FIGURE 3.1. The blue line is the integral of the cublic spline fit of ξ_0^T (initial instantaneous forward variance); the dots represent maturity times variance swap market data on August 18, 2021 of DAX Index.

To the above Λ's (2.3)-(2.7) correspond Bergomi-Guyon expansions $\mathcal{S}_{\text{2fB}}(T)$, $\mathcal{S}_{\text{rB}}(T)$, $\mathcal{S}_{\text{CPL}}(T)$, $\mathcal{S}_{\text{TSPL}}(T)$, $\mathcal{S}_{\text{PLI}}(T)$ of the ATM skew (2.2). For instance, $\mathcal{S}_{\text{2fB}}(T) = \text{B}_{k_1,k_2,c_1,c_2}(T)$ depends on the four parameters k_1, k_2, c_1, c_2 of Λ_{2fB}. Those are the five model-induced parametrizations of the term-structure of ATM skew that we consider in our comparative study. We also add the simple power-law $\text{PL}^*_{c,\alpha}(T) := cT^{-\alpha}$ and the simple capped power-law $\text{CPL}^*_{c,\alpha,\delta}(T) := c\max(T,\delta)^{-\alpha}$ parametrizations.

3. Data and Methodology

3.1. **Data.** We conduct our study for three major equity indexes: S&P 500 (SPX), Eurostoxx 50 (SX5E), and DAX. We source mid OTM SPX, SX5E, and DAX monthly implied volatilities from Bloomberg. Our data ranges from January 1, 2020 to December 31, 2021. We removed all dates where the ATM skew of the first available monthly maturity T_1 is less than 98% of the ATM skew of the second available monthly maturity T_2 ($\mathcal{S}(T_1) < 0.98\,\mathcal{S}(T_2)$); this unusual shape represents approximately 3% of traded dates. We also removed dates with 4 or less traded monthly maturities lesser than 1 year. The total number of trading dates is thus 499 for SPX, 476 for DAX, and 485 for SX5E.

Note that by mixing positive and negative spot-vol correlations, the 2fB model can actually generate non-monotonic term-structures of ATM skew and fit inverted term-structures where $\mathcal{S}(T_1) < \mathcal{S}(T_2)$ quite well. The other models and parametrizations considered here cannot do so. Thus, removing the unusual dates where $\mathcal{S}(T_1) < 0.98\,\mathcal{S}(T_2)$ actually penalizes the 2fB model and favors the other models/parametrizations.

3.2. **Methodology.** We conduct the following calibration study. For each Λ (2.3)-(2.7), on every business day in our two-year sample, and for each of the three equity indexes above,

(1) for given parameters of Λ, we numerically compute the corresponding term-structure $\mathcal{S}(T)$ using (2.2); note that it depends on the initial term-structure $u \mapsto \xi_0^u$ observed on that day;
(2) we optimize over the parameters of Λ so as to fit the observed term-structure of ATM skew as best as possible, using ordinary least squares.

We also apply (2) to find the best fitting power law $\text{PL}^*_{c,\alpha}(T)$ and capped power law $\text{CPL}^*_{c,\alpha,\delta}(T)$. To be more precise, for each business day and equity index, we do the following:

(1) **Market ATM skews.** We compute market ATM skews $\mathcal{S}(T)$ at all monthly maturities T.[2] This is done by fitting a polynomial of degree 2 locally around the money, using the 6 closest strikes to the left and the 6 closest strikes to the right of the T-forward. We have carefully checked that this local fit accurately estimates the market ATM skew, including for very short maturities.

[2]We only consider monthly maturities, as they are the most liquidly traded.

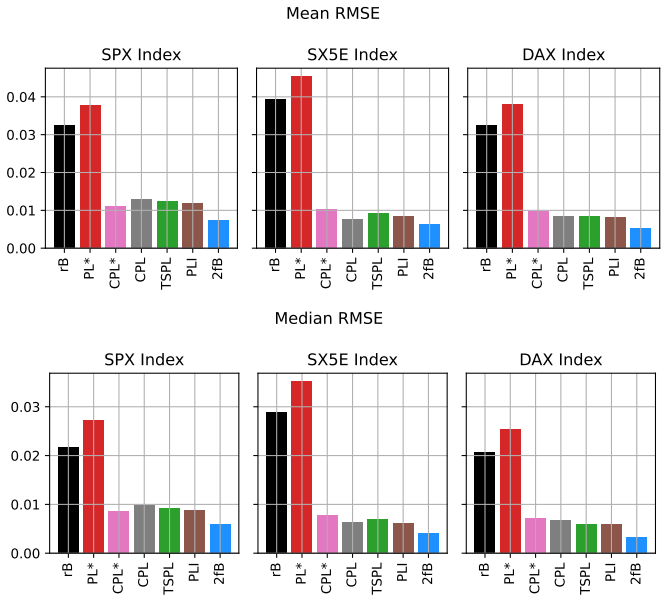

FIGURE 4.1. 0-1Y fit: mean and median RMSE comparison for SPX, SX5E, and DAX

(2) **Instantaneous forward variance.** ξ_0^u is computed from the market prices $\mathrm{VS}(T)$ of variances swaps: $\xi_0^u = \frac{d}{du}(u\mathrm{VS}(u))$. The variance swap rate of maturity T, $\mathrm{VS}(T)$ is calculated from the smile of maturity T using the replication formula from Carr and Madan [5]. The Carr-Madan formula consists of 2 integrals that we compute using the `quad` function from the `scipy` library in Python. We use market implied volatilities data of all available strikes and maturities to fit the implied volatility smile using a cubic spline. We check that this does not produce any butterfly arbitrage, and finally check that the resulting surface has no calendar spread arbitrage. In order to speed up the computations of C^μ and $C^{x\xi}$ in (2.2) we compute the values of ξ_0^u for 100 maturities u and then fit a cubic spline to them. Figure 3.1 shows an example of the resulting fit of variance swap data.

(3) **Fit.** We use a least-square minimizer to fit the parametric shapes $\mathcal{S}_{\mathrm{rB}}(T)$, $\mathcal{S}_{\mathrm{CPL}}(T)$, $\mathcal{S}_{\mathrm{TSPL}}(T)$, $\mathcal{S}_{\mathrm{PLI}}(T)$, and $\mathcal{S}_{\mathrm{2fB}}(T)$ to market data.[3] Given the global power-law-like shape of the market term-structure, we actually apply the least squares method to the log-skew, and compute the Root Mean Squared Error (RMSE) on the log-skew. We also fit the log-power-law $\log \mathrm{PL}^*_{c,\alpha}(T) = \log c - \alpha \log T$ and log-capped-power-law $\log \mathrm{CPL}^*_{c,\alpha,\delta}(T) = \log c - \alpha \log \max(T, \delta)$ curves. For each parametric shape, four fits are actually performed, depending on the range of maturities considered:
   (a) **0-1Y**: To focus on the question of whether the ATM skew blows up or not at zero maturity, we only consider short maturities, not exceeding 1 year. In particular, we will investigate how the mean RMSEs depend on the time T_1 to the first monthly maturity. The average RMSE of a good model/parametrization should hardly depend on T_1.
   (b) **0-3Y**: To fit the global shape of the term-structure, we consider all maturities not exceeding 3 years. **1M-3Y** and **2M-3Y**: Like 0-3Y, but we remove the first (resp., first two) monthly maturity.[4] By comparing the results of the 1M-3Y and 2M-3Y fits to those of the 0-3Y fit we may identify parametrizations that globally fit well the term-structure except for short maturities, such as the power law in Figure 1.1.

For each index, we then compare the RMSEs of the different parametrizations.

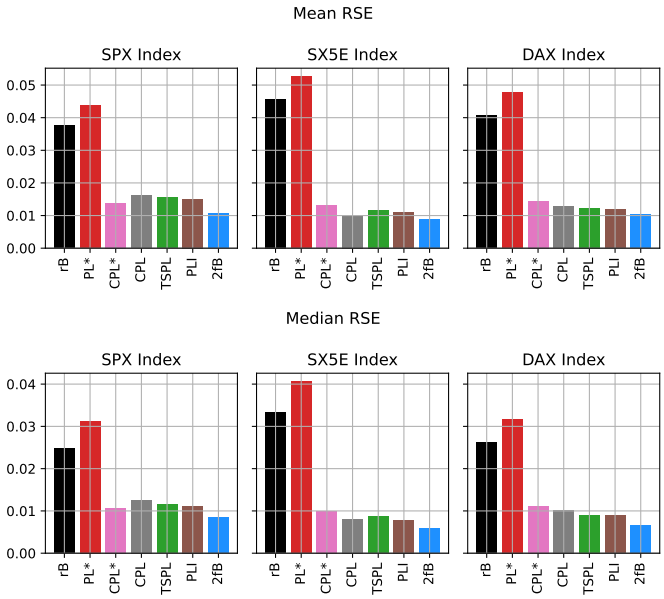

FIGURE 4.2. 0-1Y fit: mean and median RSE comparison for SPX, SX5E, and DAX

4. RESULTS

We first look at the 0-1Y fit. Figure 4.1 compares the mean RMSEs of the different parametric shapes, averaged over all calibration dates. Not surprisingly, the more parameters, the smaller the mean RMSE. However, it is clear from the figure that the increase in mean RMSE when we move from the 3-parameter shapes (CPL*, CPL, TSPL, PLI) to the 2-parameter shapes (rB, PL*) cannot be merely attributed to the number of parameters. Indeed, depending on the index, the rB mean RMSEs are 2.5 to 5.1 times larger than the CPL, TSPL, and PLI ones, while the latter are only 1.2 to 1.7 times larger than the 2fB ones. The median RMSEs, also reported in Figure 4.1, prove that those differences are not due to outlier RMSEs. The full distributions of RMSEs are reported in Appendix A. To account for the number of parameters in each model, we also compute the mean and median residual standard errors (RSEs); Figure 4.2 confirms that the large mean RMSEs of rB/PL* cannot be only attributed to the smaller number of parameters.

Since the rB/PL* shapes can very well approximate the other shapes on the 0-1Y range of maturities except for very short expiries, where rB/PL* blow up while the other shapes do not, the increase in mean RMSE when we move from CPL*/CPL/TSPL/PLI/2fB to rB/PL* is likely in large part due to the fact that the power law shapes (rB, PL*) do not have enough flexibility to accurately fit both the very short skew and longer skews, even when we only consider the limited 0-1Y range. This is a first sign that the equity ATM skew does not seem to follow a power law on the whole 0-1Y range.

To confirm this, let us look deeper into the data. In Figure 4.3 we compute the average RMSE as a function of the time T_1 to the first monthly option maturity. (The graph for the median RMSE is given in Appendix B.) The average RMSE of a good model should hardly depend on T_1. This is the case for all the 3- and 4-parameter models. That the rB and PL* average RMSEs increase a lot as T_1 gets smaller is a very clear sign that ATM skews do not blow up like the 0-1Y best power law fit for vanishing maturities. This dependency on T_1 is also reflected in the seasonality of the RMSE and optimal parameters for rB and PL* (see Figures 4.4 and 4.6). The RMSE and optimal parameters have a period of about one month. The RMSE peaks around the middle of the month, just before the monthly option expiries, and then jumps down to its minimum value immediately after. No seasonality is observed for the other shapes (see Figures 4.5 and 4.7).

[3]Equation (2.2) includes double and triple integrals. We use a Gaussian quadrature with 8 points per dimension to compute C^{μ} and $C^{x\xi}$ for each kernel. We have checked that 8 points are enough to get a very accurate result.

[4]In fact, due to the large number of quoted option tickers (especially for SPX) and data download limits, the largest fitted maturity may be below 3 years. It is on average 430 days for SPX, 991 days for SX5E, and 625 days for DAX.

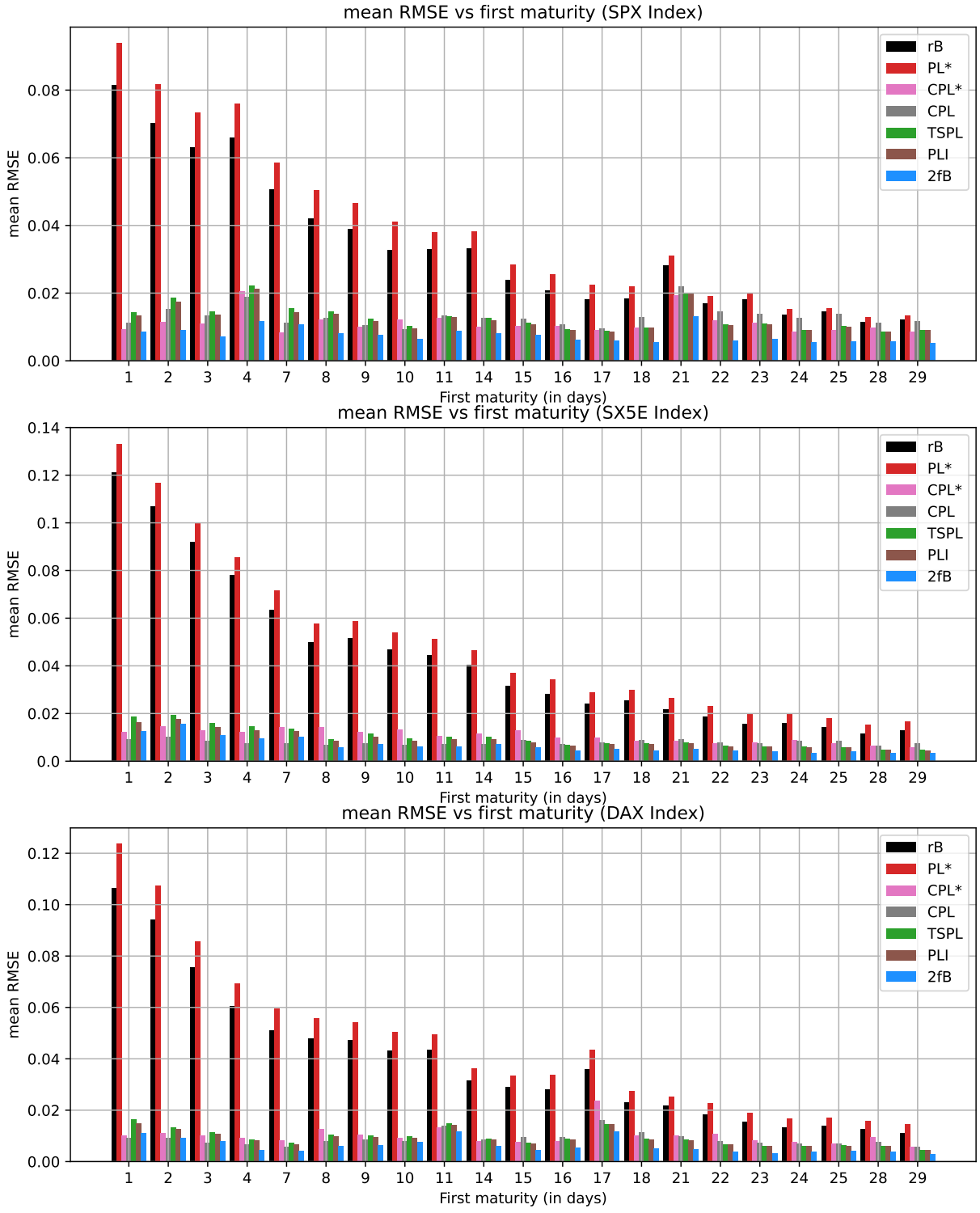

FIGURE 4.3. Mean RMSE as a function of the time T_1 to the first monthly maturity

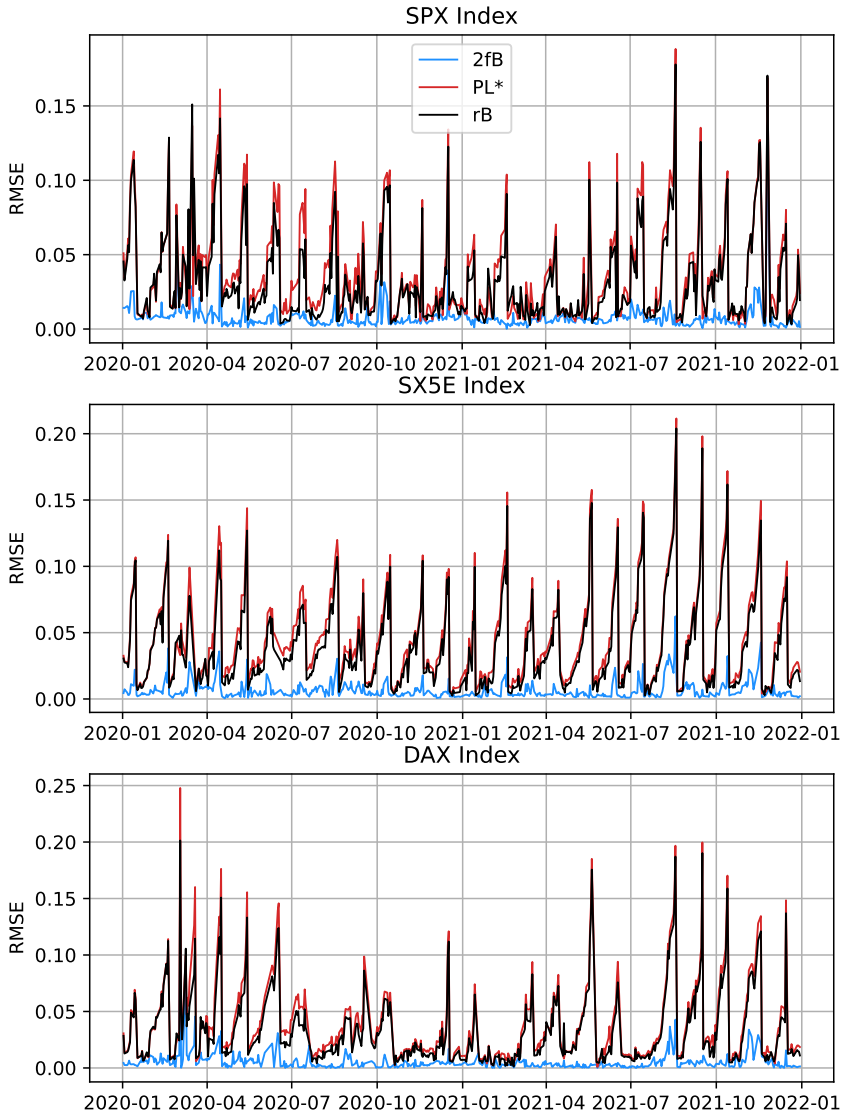

FIGURE 4.4. Time series of RMSE of 2fB, rB, and PL models for each index

The example in Figure 4.8 illustrates that rB and PL* fit the data quite well when T_1 is large; those in Figure 4.9, which are similar to the example of [12, Section 4.4], that rB and PL* fit the data quite poorly when T_1 is small, while the 3- and 4-parameter shapes fit the data well.

Let us now turn to the longer-range 0-3Y, 1M-3Y, and 2M-3Y fits. Figure 4.10 shows that, when we ignore the first or first two monthly maturities (center and right columns), rB and PL* fit the term-structure of ATM skew reasonably well over a large range of maturities. In this case, CPL*, CPL, TSPL, PLI, and 2fB fit better, but the lower median RMSE seems to be explained in large part by the larger number of parameters. By contrast, when all maturities are included, the rB and PL* fits deteriorate a lot, much more than the 3- and 4-parameter fits (see Table 1). This confirms that the global shape of the term-structure of ATM skews over a large range of maturities is quite well captured by a power law, but that the power law shape fails to accurately fit the short maturity skew. The fact that all our simple 3- and 4-parameter models fit all maturities well and have finite zero-maturity ATM skew is a sign that the ATM skew may not blow up for short maturities.

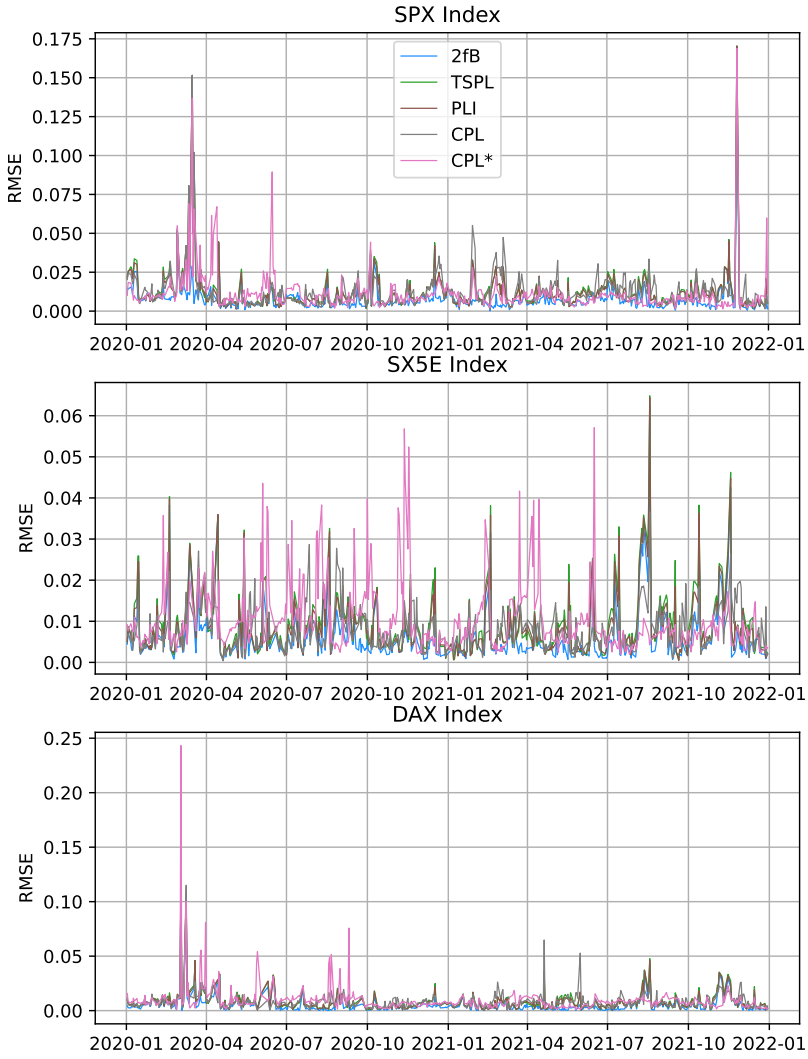

FIGURE 4.5. Time series of RMSE of 2fB, TSPL, PLI, CPL, and CPL* models for each index

		rB	PL*	CPL*	CPL	TSPL	PLI	2fB
with all maturities / without 1st maturity	**SPX**	3.04	3.46	1.41	1.39	1.54	1.51	1.40
	SX5E	2.56	2.49	1.40	1.10	1.32	1.19	1.58
	DAX	2.28	2.28	1.73	1.06	1.47	1.28	1.35
with all maturities / without 1st and 2nd maturities	**SPX**	4.25	4.98	1.75	1.84	1.97	1.90	1.64
	SX5E	3.68	3.72	2.09	1.17	1.92	1.88	1.83
	DAX	3.02	3.30	2.06	1.36	2.45	2.42	1.65

TABLE 1. Ratios of median RMSEs from Figure 4.10

Figure 4.14 shows the distributions of extrapolated ATM skews at $T = 0$ for the 3- and 4-parameter models. Note that the zero-maturity skew, far from being infinite, is distributed around 1.5. This value of 1.5 thus seems to be the correct order of magnitude for the zero-maturity skew.

The fact that the term-structure of ATM skew follows a power law for large enough maturities, but that the shape changes for short maturities, is also strongly supported by Figure 4.12. The figure shows that the average ATM skew closely follows a power law for $T \geq 3$ weeks, but that it is almost capped flat for $T \leq 3$

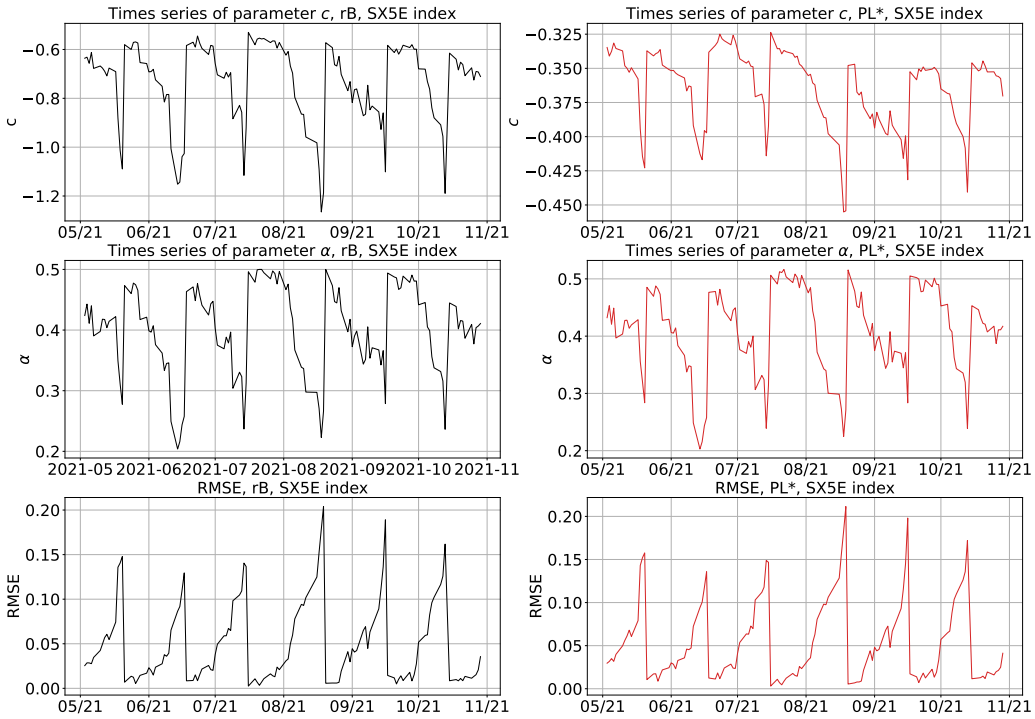

FIGURE 4.6. Time series of rB and PL optimal parameters and RMSE between May and November 2021 on SX5E index.

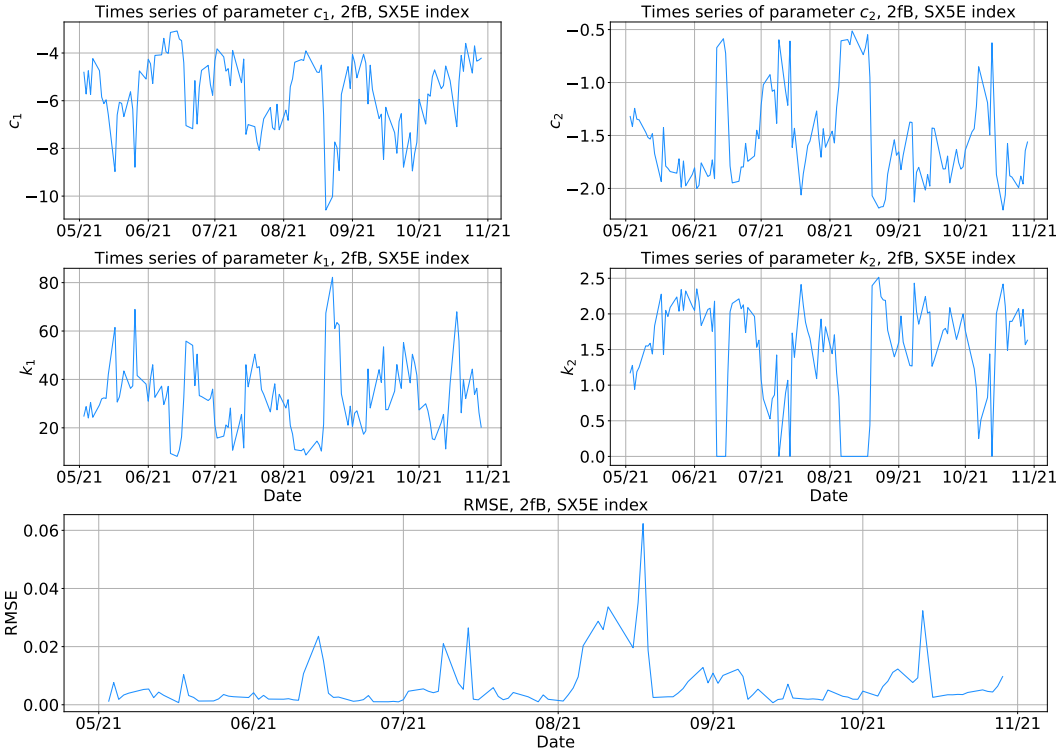

FIGURE 4.7. Time series of 2fB optimal parameters and RMSE between May and November 2021 on SX5E index.

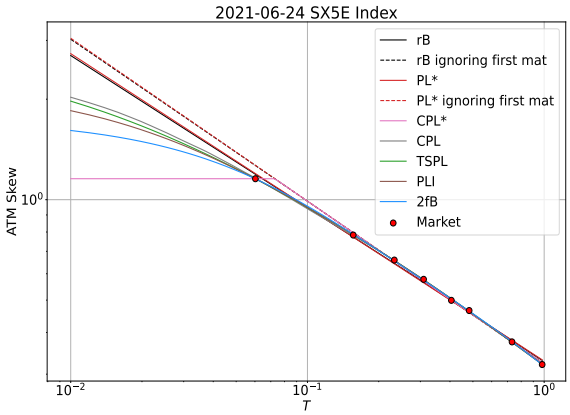

FIGURE 4.8. SX5E ATM skew as of June 24, 2021. Comparison of the fits of different parametric shapes.

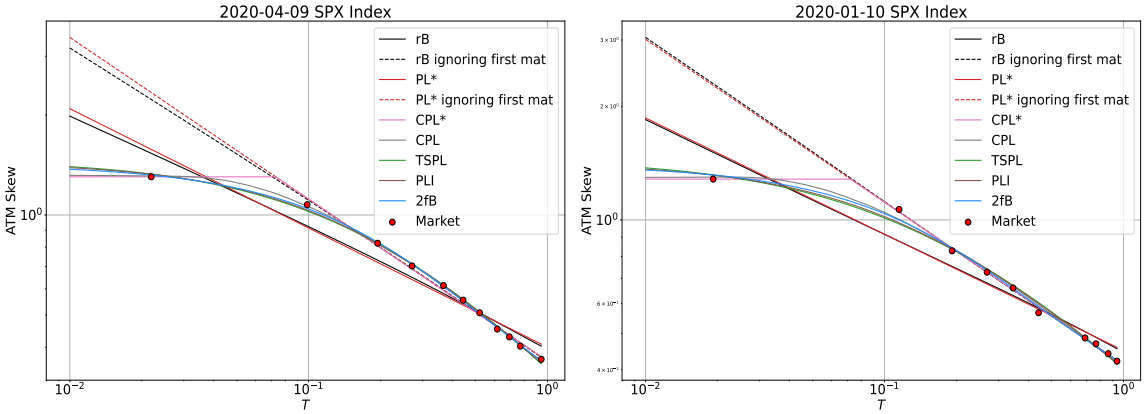

FIGURE 4.9. SPX ATM skew as of January 10, 2020 (left) and April 9, 2020 (right). Comparison of the fits of different parametric shapes.

	SPX	**SX5E**	**DAX**
For maturities below 3 weeks	0.19	0.04	0.08
For maturities above 3 weeks	0.43	0.44	0.45

TABLE 2. Power law exponents α for maturities below 3 weeks and maturities above 3 weeks, based on Figure 4.12

weeks (for SX5E and DAX) or follows another, much "flatter" power law for $T \leq 3$ weeks (for SPX). Table 2 gives the corresponding power law exponents α. The average ATM skew of SX5E and DAX does not seem to blow up at zero maturity. It is remarkable that for the three indexes the average term-structure has such a clear kink at the critical maturity $T^* = 3$ weeks. This critical maturity also shows up when we study individual term-structures. Indeed, Figure 4.11 shows that when the δ parameter of CPL* is smaller than T_1 (in which case CPL* simply acts as the uncapped power law PL*), then T_1 is most often larger than 3 weeks; and when $\delta > T_1$ (in which case the cap is used to improve the power law fit), then δ is distributed around an average of 3 weeks to 1 month.

Our results suggest that if one needs to use a pricing model that correctly captures the ATM skew over, say, the full 0-1Y or 0-3Y maturity range, including short maturities below 1 month, one should use the CPL, TSPL, or PLI versions of Model (2.9) or the 2fB model rather than models producing a power-law term structure, such as rough volatility models. The 2fB model is much more handy than the CPL, TSPL,

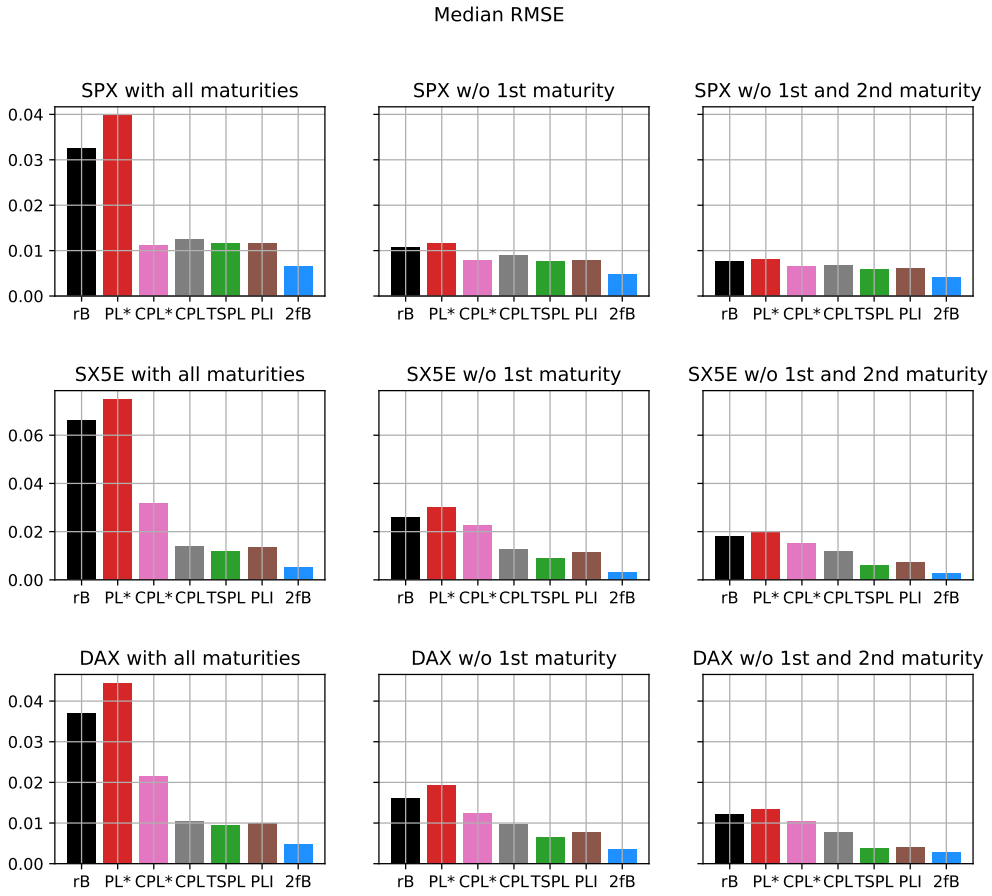

FIGURE 4.10. 0-3Y, 1M-3Y, and 2M-3Y fits: median RMSE comparison for SPX, SX5E, and DAX

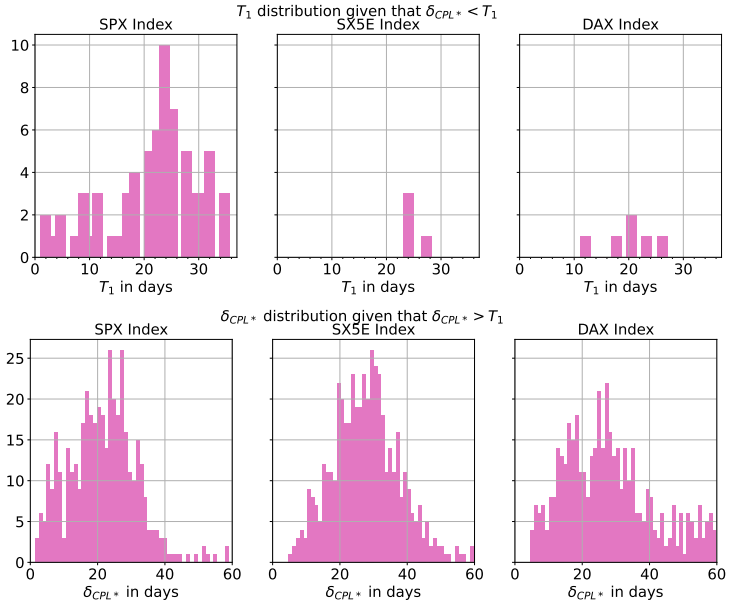

FIGURE 4.11. Conditional distributions of T_1 and δ for CPL*

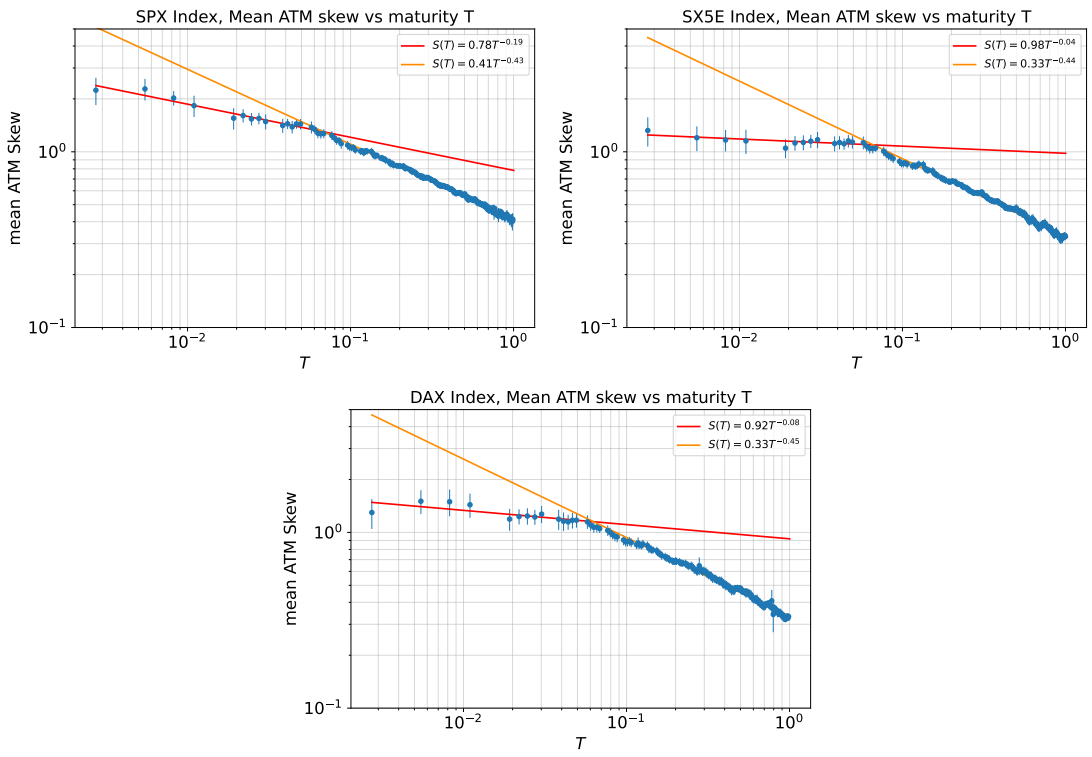

FIGURE 4.12. Average term-structures of ATM skew in log-log scale over two years of data (2020 and 2021) for SPX, SX5E, and DAX

PLI, and rough volatility models: it is Markovian, thus very easy and quick to simulate, whereas the latter are not. Note that since the ATM skew only depends on the spot-vol correlation $\Lambda = \lambda_1$, the number d of Brownian motions driving the whole volatility dynamics (2.1) is irrelevant to fitting the skew. In particular, one may degenerate the classical 2fB model into a one-factor model, or even into a zero-factor model, i.e., a pure path-dependent volatility (PDV) model, similar to the 4-factor[5] Markovian PDV model of Guyon and Lekeufack [14].

While the Bergomi-Guyon expansion only provides an approximation (at order 2 in the volatility-of-volatility) of a model ATM skew, we check in Figure 4.13, in the case of the 2fB model, that this approximation is quite accurate, even for very large volatility-of-volatility ($\omega = 8.5$). We use average model parameters from Figure 4.7. Figure 4.13 also shows that, in this example, the expansion almost perfectly fits the true term-structure of ATM skew of a 2fB model: the original model with a slightly decreased volatility-of-volatility $\omega \to \omega'$.

5. CONCLUSION

Our empirical study confirms that a power law fits the term-structure of ATM skew of equity indexes quite well over a large range of maturities, from 1 month to a few years. But this should not lead one to conclude that the ATM skew blows up like this power law at zero maturity. Zooming in on the short maturities shows that the power-law extrapolation fits the data poorly. We have shown that simple models, which generate finite zero-maturity skews, fit the whole term-structure—including short maturities below 1 month—much better than the rough Bergomi model or a power law, whose goodness of fit deteriorate quickly as the first monthly option maturity gets closer. The average term-structures indicate a critical maturity of 3 weeks. Our study concludes that, far from being infinite, the zero-maturity extrapolation of the skew is distributed around a typical value of 1.5 (in absolute value).

[5]Here "4-factor" means that 4 jointly Markovian variables are sufficient to compute the volatility.

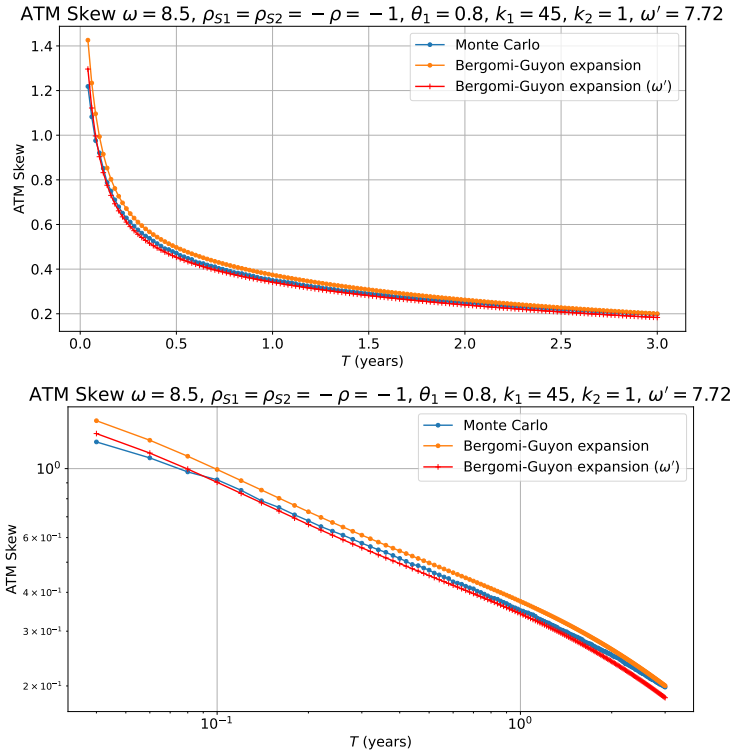

FIGURE 4.13. Term-structure of the ATM skew in the 2fB model: Monte Carlo vs Bergomi-Guyon expansion. Top: linear scale. Bottom: log-log scale

Acknowledgements. We would like to thank the anonymous referees for their constructive comments and suggestions that helped improve a preliminary version of this article.

REFERENCES

[1] Alòs, E., León, J.A., Vives, J.: *On the short-time behavior of the implied volatility for jump-diffusion models with stochastic volatility*, Finance Stoch. 11:571–589, 2007.
[2] Bayer, C., Friz, P., Gatheral, J.: *Pricing under rough volatility*, Quantitative Finance, 16(6):887–904, 2016.
[3] Bergomi, L.: *Smile dynamics II*, Risk, October 2005.
[4] Bergomi, L., Guyon, J.: *Stochastic volatility's orderly smiles*, Risk, May 2012.
[5] Carr, P., Madan, D.:*Option valuation using the fast Fourier transform*, Journal of Computational Finance 2(4):61–73, 1999.
[6] Fouque, J.-P., Papanicolaou, G., Sircar, R., Solna, K.: *Maturity cycles in implied volatility*, Finance Stoch. 8:451–477, 2004.
[7] Fukasawa, M.: *Asymptotic analysis for stochastic volatility: Martingale expansion*, Finance Stoch. 15(4):635–654, 2011.
[8] Fukasawa, M.: *Volatility has to be rough*, Quantitative Finance 21(1):1–8, 2021.
[9] Gatheral, J.: *Consistent modeling of SPX and VIX options*, presentation at Bachelier Congress, July 18, 2008.
[10] Gatheral, J., Jaisson, T., Rosenbaum, M.: *Volatility is rough*, Quantitative Finance 18(6):933–949, 2018.
[11] Gatheral, J., Kamal, M.: *Implied Volatility Surface*, in Encyclopedia of Quantitative Finance, John Wiley and Sons, 2010.
[12] Guyon, J.: *The VIX Future in Bergomi Models: Fast Approximation Formulas and Joint Calibration with S&P 500 Skew*, SIAM Journal on Financial Mathematics 13(4):1418–1485, 2022.
[13] Guyon, J.: *The smile of stochastic volatility: Revisiting the Bergomi-Guyon expansion*, preprint, available at `ssrn.com/abstract=3956786`, 2021.
[14] Guyon, J., Lekeufack, J.: *Volatility is (mostly) path-dependent*, preprint, available at `ssrn.com/abstract=4174589`, 2022.
[15] Rømer, S.E.: *Empirical analysis of rough and classical stochastic volatility models to the SPX and VIX markets*, Quantitative Finance 22(10):1805–1838, 2022.

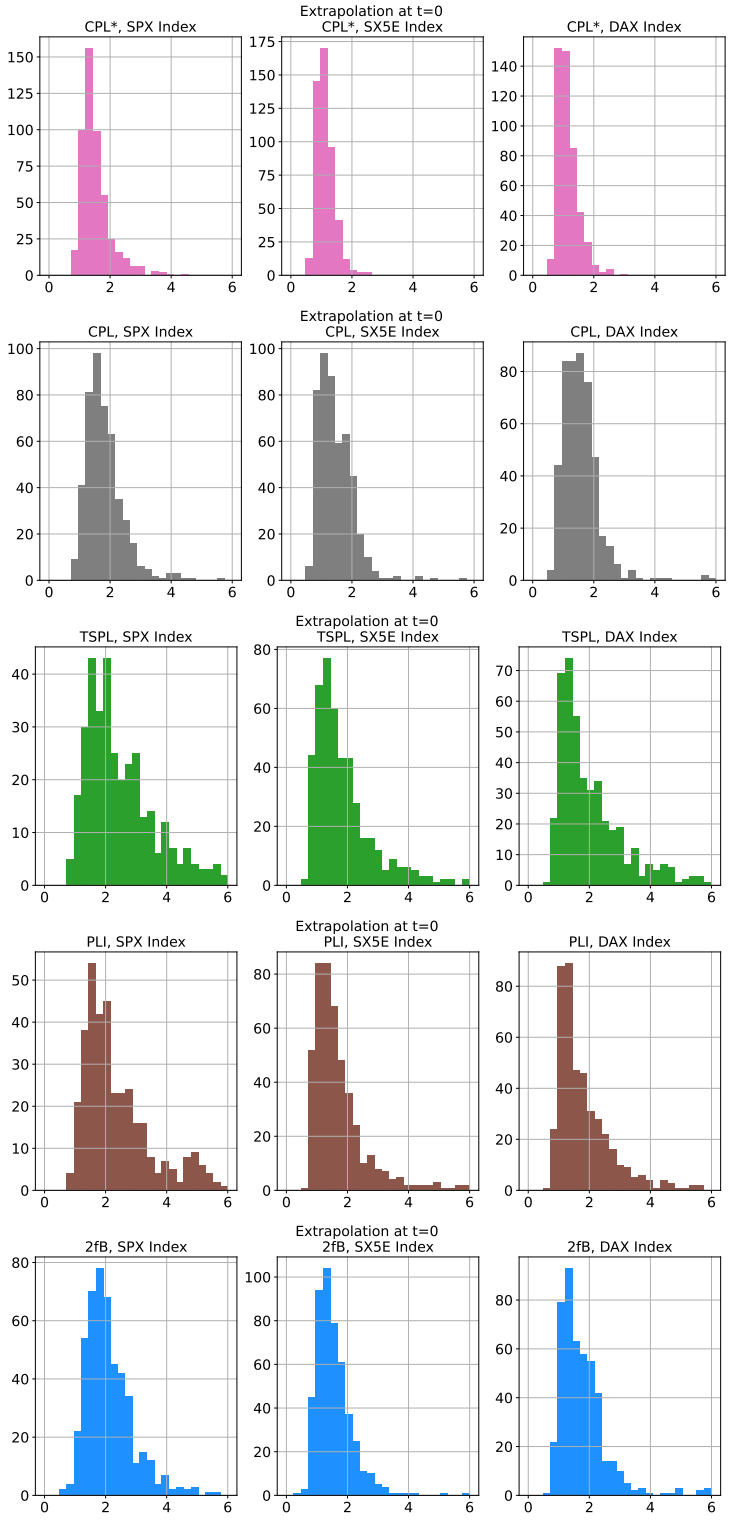

FIGURE 4.14. Extrapolation at $T = 0$ of the ATM skew for SPX, SX5E, and DAX. From top to bottom: CPL*, CPL, TSPL, PLI, and 2fB

Appendices

Appendix A. Histograms of RMSEs

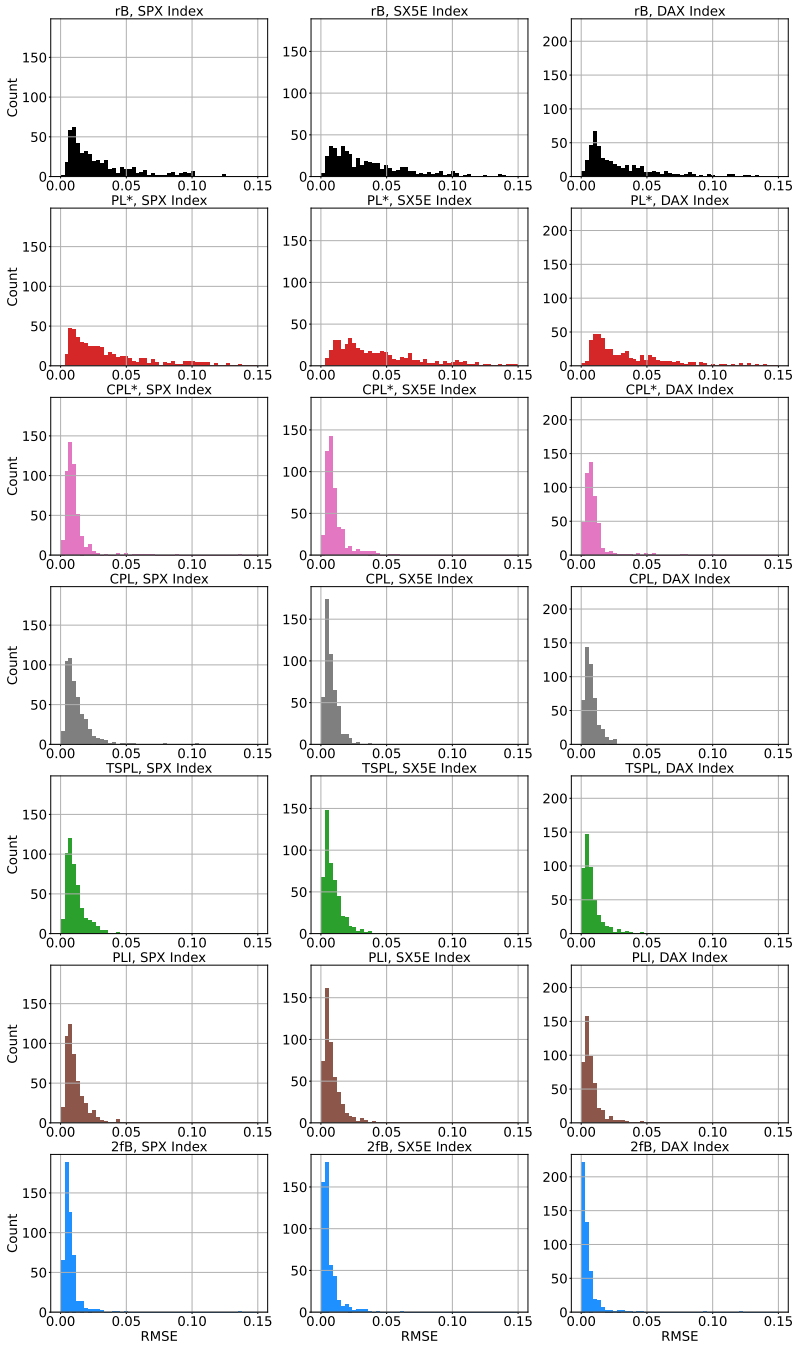

Figure A.1. 0-1Y fit: histogram of RMSE per parametric shape (rows) per index (columns)

Appendix B. Median RMSE as a function of T_1

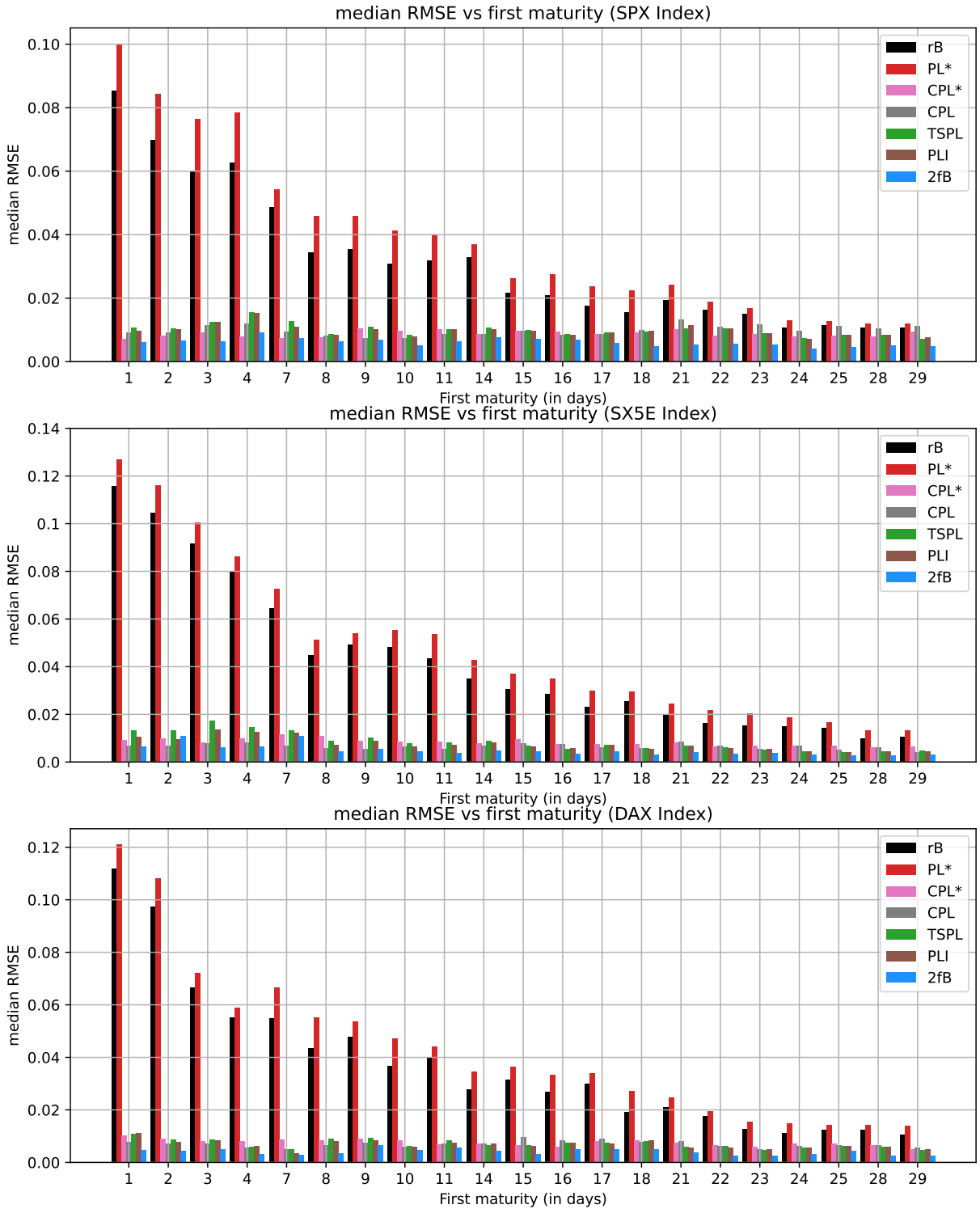

Figure B.1. Median RMSE as a function of the time T_1 to the first monthly maturity