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

MEHDI EL AMRANI AND JULIEN GUYON

Bloomberg L.P., Quantitative Research

Abstract. Using two years of S&P 500, Eurostoxx 50, and DAX data, we empirically show that the term-structure of the at-the-money (ATM) skew of equity indexes does *not* follow a power law for short maturities and is much better captured by simple parametrizations that do *not* blow up for vanishing maturity. The ATM skew produced bythe two-factor Bergomi model provides the best fits. Two other models built using non-blowing-up kernels are introduced and shown to produce similar results. By contrast, the fits of the rough Bergomi model and power law deteriorate quickly as we get closer to the first monthly options maturity. The extrapolated zero-maturity skew, far from being infinite, is distributed around 1.5 (in absolute value).

Keywords. At-the-money skew, equity skew, variance curve models, calibration.

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 (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 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 to? 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 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 two-factor Bergomi 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 add 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 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 conclude; the main contribution of this article is precisely to perform a systematic empirical study covering several 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

E-mail address: melamranizi1@bloomberg.net, jguyon2@bloomberg.net.

Date: First version: February 17, 2022. This version: July 25, 2022.

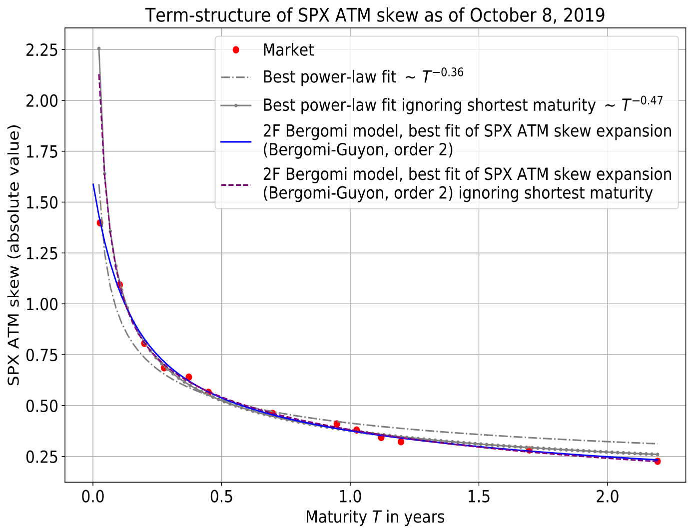

FIGURE 1.1. SPX ATM skew as of October 8, 2019. Comparison of power-law fits and fits using the second-order Bergomi-Guyon expansion of the SPX ATM skew in the two-factor Bergomi model

ATM skews of equity indexes blow up for very short maturities, like in rough volatility models. Our empirical study shows that, *consistently over time and across equity indexes*, the term-structure of ATM skew is actually better captured by parametric shapes which *do not blow up* for vanishing maturities: $\mathrm{B}_{k_1,k_2,c_1,c_2}(T)$, as well as other parametric shapes (TSPL and PLI) which depend on only three parameters—just one more than the power law; see Figures 3.2, 4.2, and 5.1. Those parametric shapes are produced by stochastic volatility models in which the ATM skew does not blow up for vanishing maturities. We thus argue that such blow-up should not be taken for granted. Not only is it not verifiable on market data; it is also inconsistent with simple models that fit market data better—not surprisingly, since they have a bit more parameters. Among the models that we test, the smallest root mean squared error (RMSE) scores are consistently obtained by the two-factor Bergomi model. (Note that, being Markovian, this model is also much more handy than rough volatility models.) Our study shows in particular that the two-factor Bergomi model fits the market term-structure better than the power law $\mathrm{PL}_{c,\alpha}(T) := cT^{-\alpha}$ or the rough Bergomi model on more than 99% of the trading days in 2020 and 2021, consistently across equity indexes.

We shall also see that the RMSE scores of the non-blowing-up models hardly depend on the time to the first monthly option maturity, while the RMSE scores of the rough Bergomi model and pure power-law decay deteriorate significantly as the time to the first monthly option maturity goes to zero (see Figures 4.5 and 4.6). This provides strong additional evidence that the ATM skew of equity indices actually does not follow a power law for short maturities and does not blow up for vanishing maturities.

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

$$\frac{dS_t}{S_t} = (r_t - q_t)\,dt + \sqrt{\xi_t^t}\,dW_t^1 \tag{2.1}$$

$$d\xi_t^u = \lambda(t,u,\xi_t^\cdot)\cdot dW_t, \qquad u>0, \quad t\in[0,u) \tag{2.2}$$

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

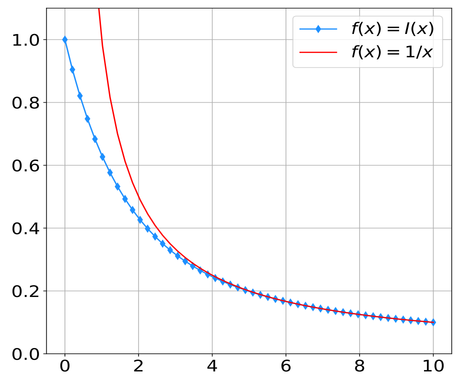

FIGURE 2.1. Graph of the functions I and $x \mapsto 1/x$

function, which encapsulates all the spot-vol covariance information:

$$\mathcal{S}(T) = \sigma_{\mathrm{VS}} \left(\frac{C^{x\xi}}{2v^2} + \frac{4C^{\mu}v - 3(C^{x\xi})^2}{8v^3} \right) + O(\omega^3), \tag{2.3}$$

where $v := \int_0^T \xi_0^u \, du$ denotes the integrated variance over $[0,T]$ seen from time 0, $\sigma_{\mathrm{VS}} := \sqrt{v/T}$ is the corresponding variance swap rate, and

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

For example, for the two-factor Bergomi model, Λ is of the form

$$\Lambda_{2\mathrm{fB}}(t,u,y) = \left(c_1 e^{-k_1(u-t)} + c_2 e^{-k_2(u-t)} \right) y^u, \qquad c_1, c_2 \in \mathbb{R}, \quad k_1, k_2 > 0 \tag{2.4}$$

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

$$\text{rough Bergomi:} \qquad \Lambda_{\mathrm{rB}}(t,u,y) := c(u-t)^{-\alpha} y^u, \qquad c \in \mathbb{R}, \quad \alpha \in (0,1/2) \tag{2.5}$$

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

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

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)

$$\frac{d\xi_t^u}{\xi_t^u} = K(u-t)(\rho\, dW_t^1 + \sqrt{1-\rho^2}\, dW_t^2) \tag{2.9}$$

with the convolution kernel K being one of the following:

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

K_{rB} is the rough Bergomi (rB) kernel [2]; K_{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. Those last two kernels (or models, or Λs) have been suggested in [13, Section 4.3] precisely in order to reproduce the power-law like decay (away from ultrashort maturities) of the term-structure of ATM skews, including for very large maturities, without generating a blow-up for vanishing maturities. They have only one extra parameter (3 in total) compared to the rough Bergomi Λ (2 parameters); and one less parameter than the two-factor Bergomi Λ (4 parameters).

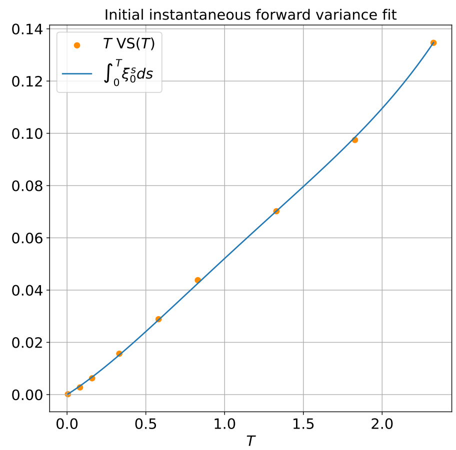

FIGURE 3.1. The blue line is the integral of the cublic spline fit of ξ_0^T (initial instantaneous forward variance); the dots represent maturity times variance swap market data on August 18, 2021 of DAX Index.

Note that, in contrast with the two-factor Bergomi model, the three kernels (2.10) produce non-Markovian models.

To the above Λ's (2.4)-(2.7) correspond Bergomi-Guyon expansions $\mathcal{S}_{\mathrm{2fB}}(T)$, $\mathcal{S}_{\mathrm{rB}}(T)$, $\mathcal{S}_{\mathrm{TSPL}}(T)$, $\mathcal{S}_{\mathrm{PLI}}(T)$ of the ATM skew (2.3). For instance, $\mathcal{S}_{\mathrm{2fB}}(T) = \mathrm{B}_{k_1,k_2,c_1,c_2}(T)$ depends on the four parameters k_1, k_2, c_1, c_2 of Λ_{2fB}. Those are the 4 model-induced parametrizations of the term-structure of ATM skew that we consider in our comparative study. We also add the simple power law parametrization $\mathrm{PL}_{c,\alpha}(T) := cT^{-\alpha}$, which is close to $\mathcal{S}_{\mathrm{rB}}(T)$ but ignores the initial term-structure ξ_0^t of instantaneous forward variances as well as the second-order term in the Bergomi-Guyon expansion.

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 the ATM skew of the second available monthly maturity T_2 ($\mathcal{S}(T_1) < \mathcal{S}(T_2)$). This unusual shape represents approximately 4% of traded dates. The total number of trading dates is thus 495 for SPX, 478 for DAX and 476 for SX5E. The SPX options market was particularly volatile during those two years which resulted in an increased instability of optimal parameters and errors, compared with SX5E and DAX.

Note that by mixing positive and negative spot-vol correlations, the 2-factor Bergomi 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, see Figure 4.10. The other models and parametrizations considered here cannot do so. Thus, removing the unusual dates where $\mathcal{S}(T_1) < \mathcal{S}(T_2)$ actually penalizes the 2-factor Bergomi model and favors the other models/parametrizations.

3.2. **Methodology.** We conduct the following calibration study. For each Λ (2.4)-(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 $\mathcal{S}(T)$; 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 $\mathrm{PL}_{c,\alpha}(T)$. To be more precise, for each business day t and equity index, we do the following:

(1) For each monthly maturity T not exceeding 3 years, we use market implied volatilities data of all available strikes for that maturity to fit the implied volatility smile using a cubic spline. We check

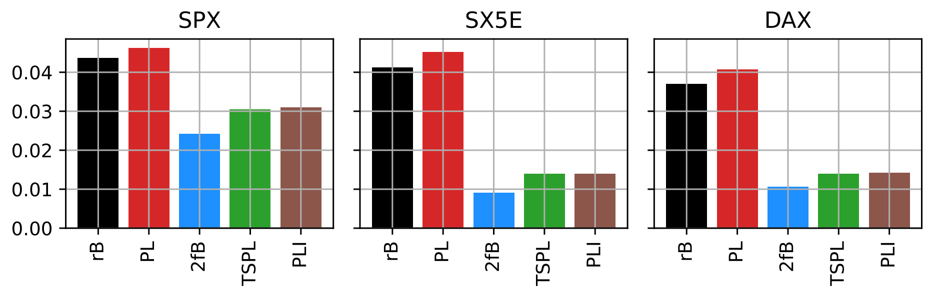

FIGURE 3.2. Mean RMSE comparison for DAX, SPX and SX5E index

that this does not produce any butterfly arbitrage. We then check that the resulting surface has no calendar spread arbitrage.

(2) We compute ATM skews at all monthly maturities.

(3) ξ_0^u is the instantaneous forward variance at time u seen from time 0. It 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 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. In order to speed up the computations of C^μ and $C^{x\xi}$ in (2.3) 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.

(4) Equation (2.3) 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.)

(5) For each Λ (2.4)-(2.7), we use a least-square minimizer to fit Equation (2.3) to market data. We add an L^2 penalization term to control the magnitude of the parameters. We obtain the optimal penalty parameters by cross-validation: $p_{\mathrm{2fB}} = 10^{-7}$, $p_{\mathrm{TSPL}} = 5 \times 10^{-9}$, $p_{\mathrm{PLI}} = 10^{-10}$.

(6) We also fit the power-law curve $\mathrm{PL}_{c,\alpha}(T)$.

(7) We compute the Root Mean Squared Error (RMSE).

We then compare the RMSEs of the different models/parametrizations, as well as the average RMSE as a function of the time to the first monthly option maturity.

4. RESULTS

The two-factor Bergomi model (2fB) performs better (smaller RMSE) than the rough Bergomi (rB) model and the power law (PL) more than 99% of the time. Figure 3.2 demonstrates that, among all the models/parametrizations tested, the two-factor Bergomi model has the lowest RMSE (averaged over calibration dates) across equity indexes: this model better captures the global term-structure of equity ATM skews. For SX5E and DAX, the RMSE of the 2fB is around 4 times smaller than the RMSE of rB and PL; for SPX, whose options market experienced a quite chaotic two years in 2020 and 2021, the RMSE of the 2fB is around 2 times smaller than the RMSE of rB and PL. The distribution of RMSE is also the narrowest and most peaked around 0 for the two-factor Bergomi model, see Figure 4.1. However, TSPL and PLI perform almost as well as 2fB. The common feature of those three models is that they produce a finite ATM skew at zero maturity. Figure 4.2 shows the time series of RMSE per index for 2fB, rB, and PL. Often, similar to the example of [12, Section 4.4], the power-law decay has not enough flexibility to accurately fit both the short and long ATM skews; see the examples in Figures 4.7 and 4.8.

Exploring the data a little bit deeper, we note the seasonality in the optimal parameters and RMSE for rB and PL (Figure 4.3). The RMSE has a period of about one month. It peaks around the end of the second week/start of the third week of the month, that is, just before the monthly expiries of the options, and then jumps to its minimum value right after. This is confirmed in Figure 4.5 where we observe that rB and PL errors increase a lot as the time T_1 to the first monthly maturity goes to 0. By contrast, the 2fB errors do not depend on T_1, see Figures 4.4 and 4.5; the TSPL and PLI errors as well, see Figure 4.6. This is a clear sign that ATM skews do *not* blow up for vanishing maturities: if it were the case, rB and PL, which produce such blow-up, would not struggle to fit the global term structure of ATM skews when T_1 gets small. On the other hand, when T_1 is large, all models tend to fit market skews quite well (even though 2fB, TSPL, and PLI still overperform rB and PL in that case). Some generate a finite zero-maturity skew (2fB, TSPL, PLI), others generate an infinite zero-maturity skew (rB and PL), and therefore even in this case one cannot conclude that the ATM skew blows up for vanishing maturities; see an example in Figure 4.9. (Actually, even in that case, one would rather conclude that the ATM skew does not blow up, since 2fB, TSPL, and PLI have a smaller RMSE.)

The ATM skew in the two-factor Bergomi model does not blow up when T vanishes. Figure 4.11 shows the distribution of extrapolated ATM skews at $T = 0$. Note that the zero-maturity skew, far from being infinite, is distributed like a lognormal density, centered around 1.5 (between 1.3 and 1.8, depending on the index). Similar graphs for TSPL and PLI are reported in Figure 4.12. They also show a distribution of the zero-maturity skew that peaks around 1.5. This value of 1.5 thus seems to be the correct order of magnitude for the zero-maturity skew.

5. Conclusion

In this article, we have empirically shown that, contrary to a common belief, the ATM skew of equity indexes does *not* follow a power law for short maturities and in particular does *not* blow up for vanishing maturities. When the time T_1 to the first monthly options maturity is large, the power law extrapolation is just one of many possible extrapolations. However, even in this case, parametrizations deriving from the two-factor Bergomi model or from other variance curve models built using non-blowing-up kernels, such as the time-shifted power-law kernel or the PLI kernel, which generate *finite* zero-maturity skews, fit the market term-structure of ATM skew better. When T_1 approaches zero, the fits of the rough Bergomi model and of the power law quickly deteriorate, a clear sign that the ATM skew does not follow such power law for short maturities. Globally, across the different values of T_1 and across equity indexes over a (challenging) data set of two years (2020 and 2021), the two-factor Bergomi model fits the term-structure of ATM skew better than the rough Bergomi model and the power law more than 99% of the time. Our empirical 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).

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*, to appear in SIAM Journal on Financial Mathematics, 2021.
[13] Guyon, J.: *The smile of stochastic volatility: Revisiting the Bergomi-Guyon expansion*, preprint, available at `ssrn.com/abstract=3956786`, 2021.

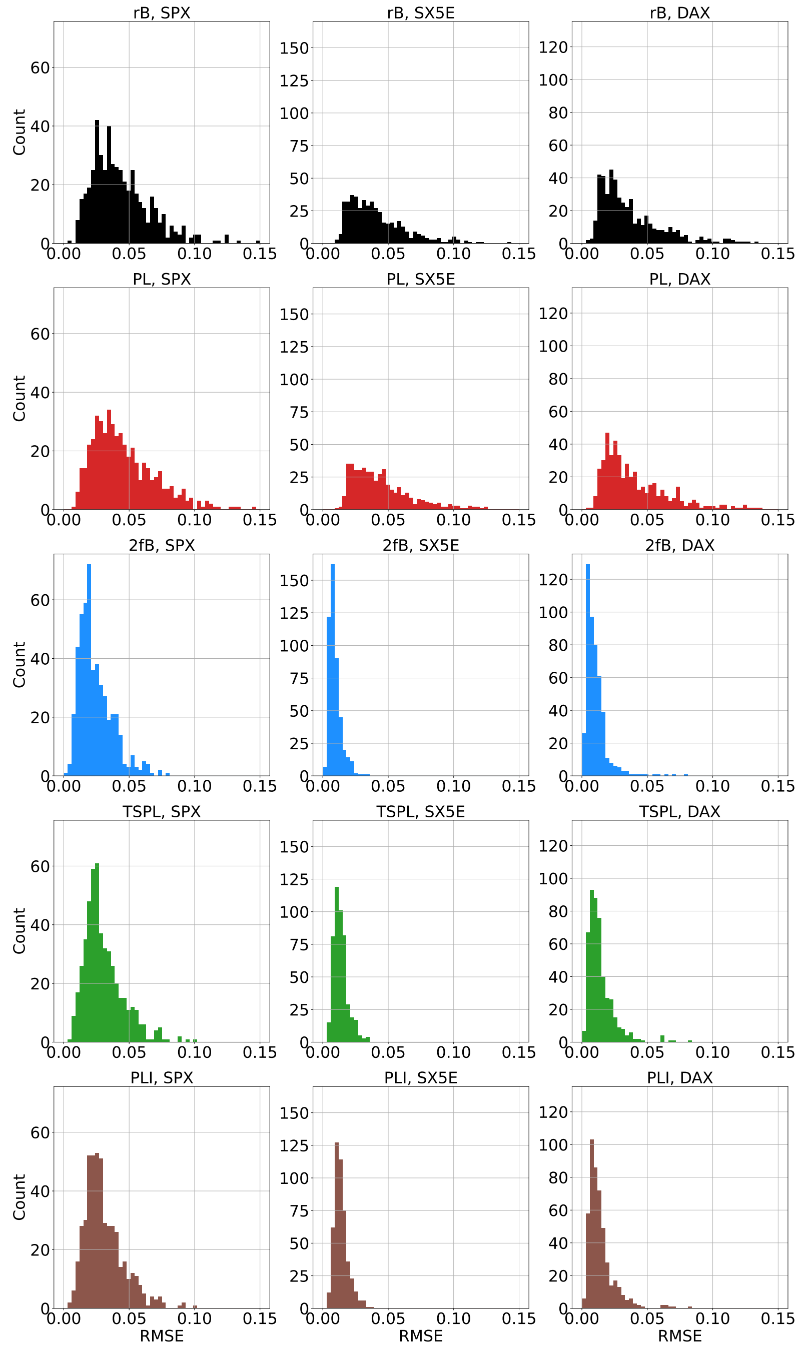

FIGURE 4.1. Histogram of RMSE per kernel (rows) per index (columns)

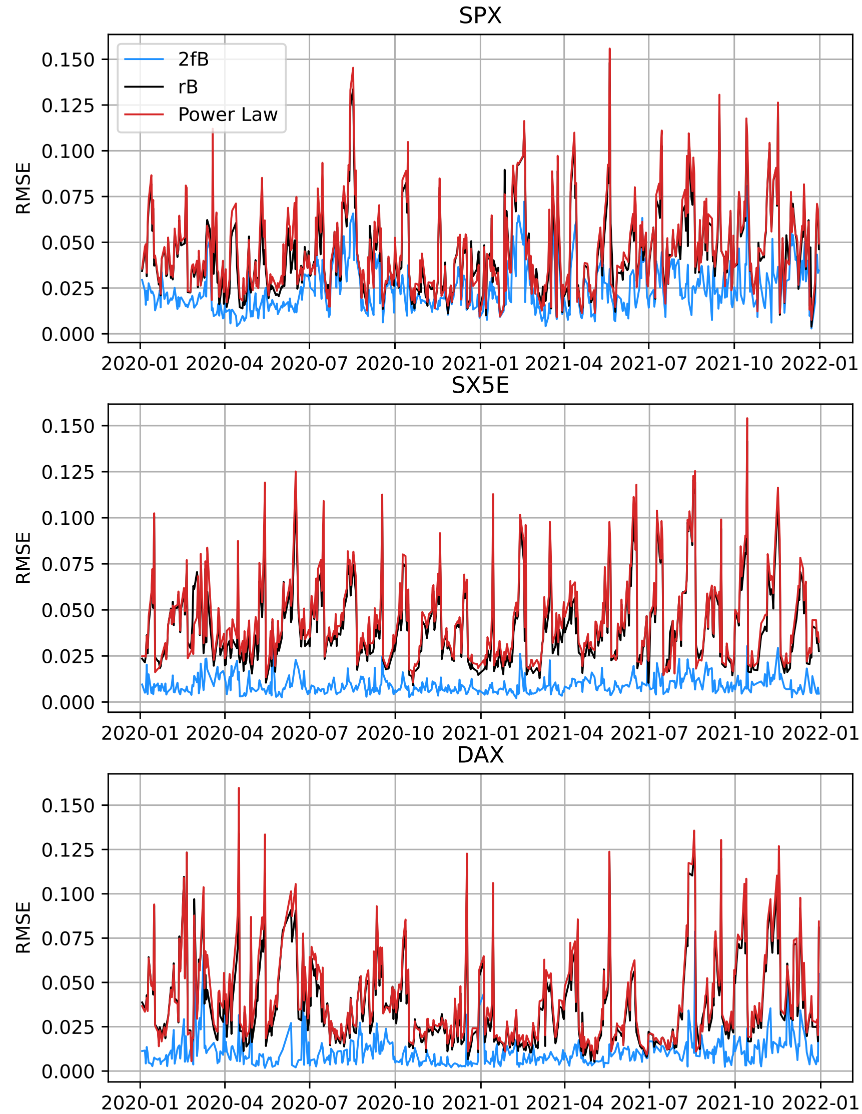

FIGURE 4.2. Time series of RMSE of 2fB, rB and PL models for each index

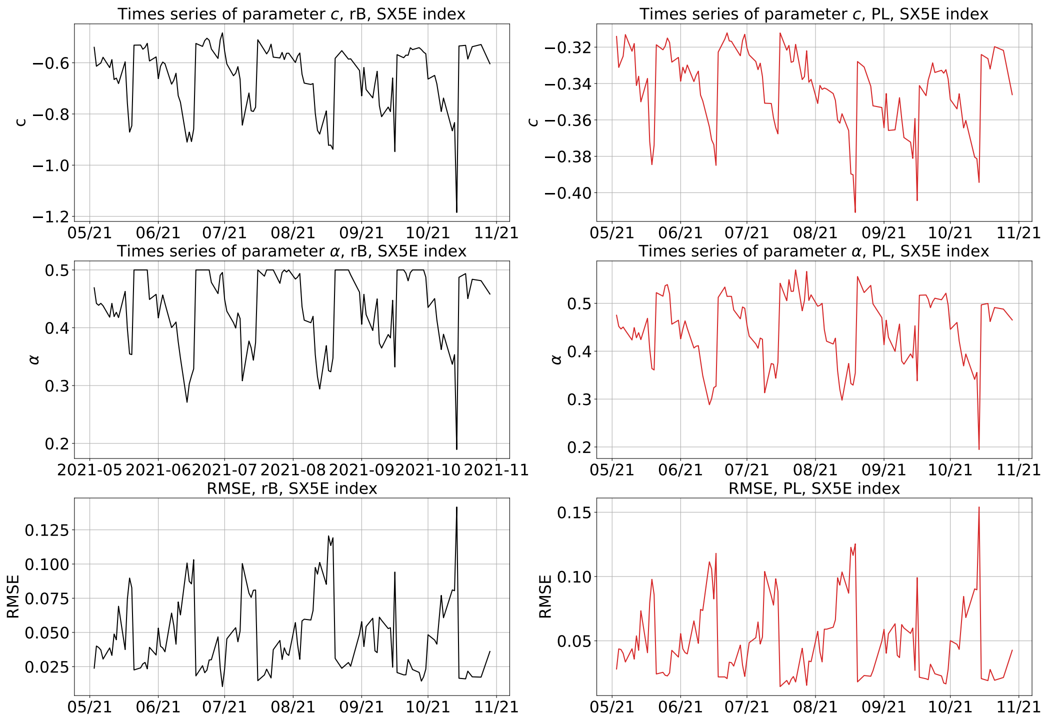

FIGURE 4.3. Time series of rB and PL optimal parameters and RMSE between May and November 2021 on SX5E index.

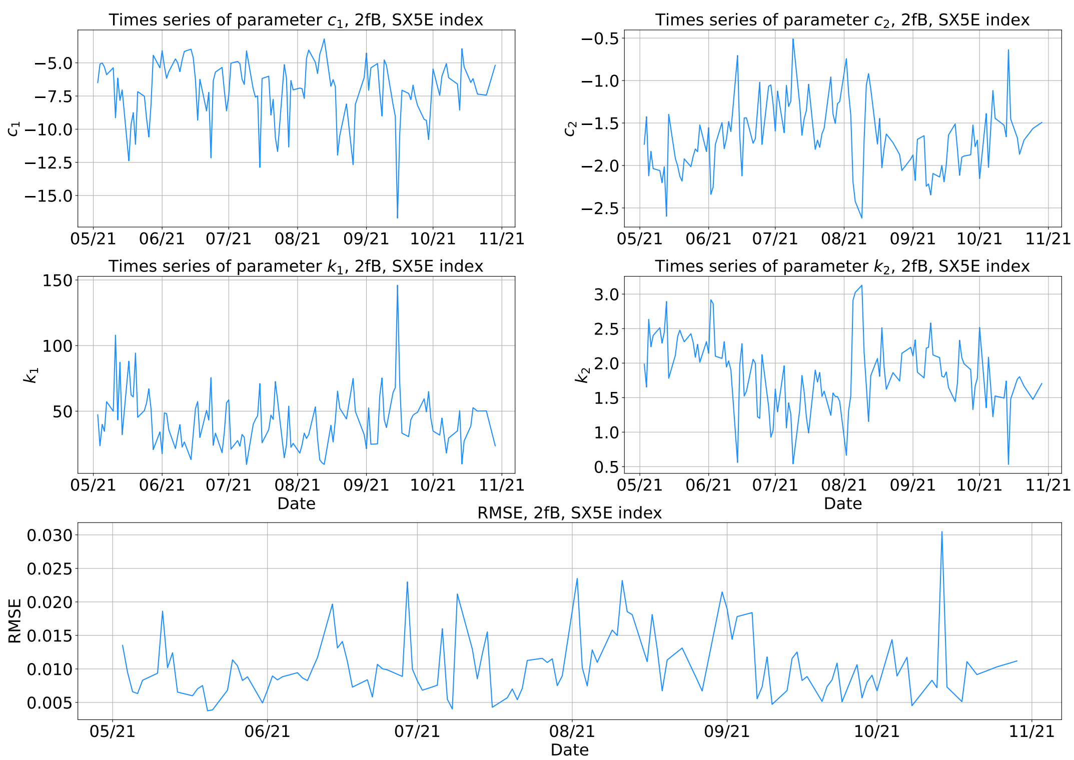

FIGURE 4.4. Time series of 2fB optimal parameters and RMSE between May and November 2021 on SX5E index.

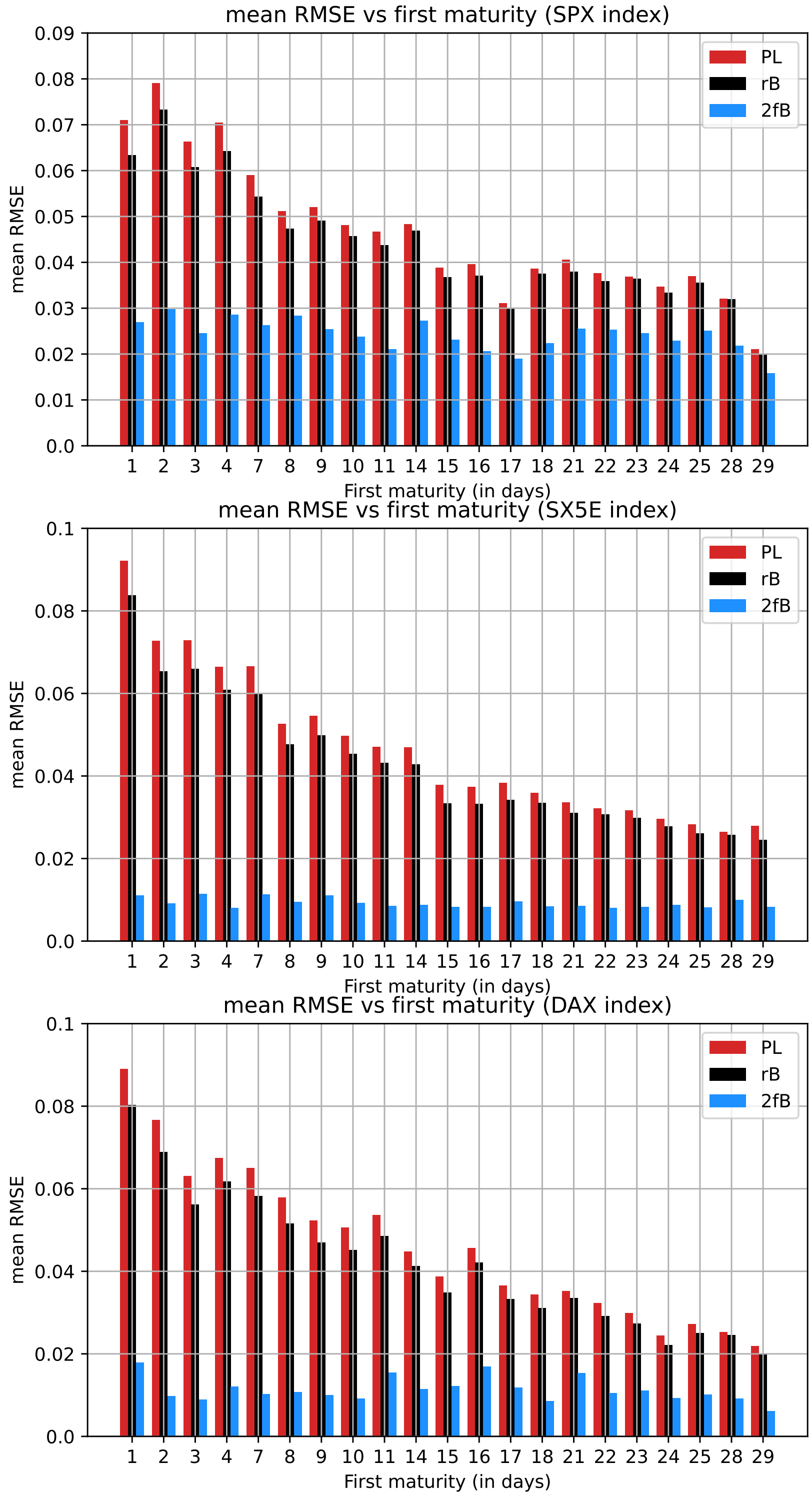

FIGURE 4.5. Mean RMSE as a function of the time to the first monthly maturity (PL, rB and 2fB)

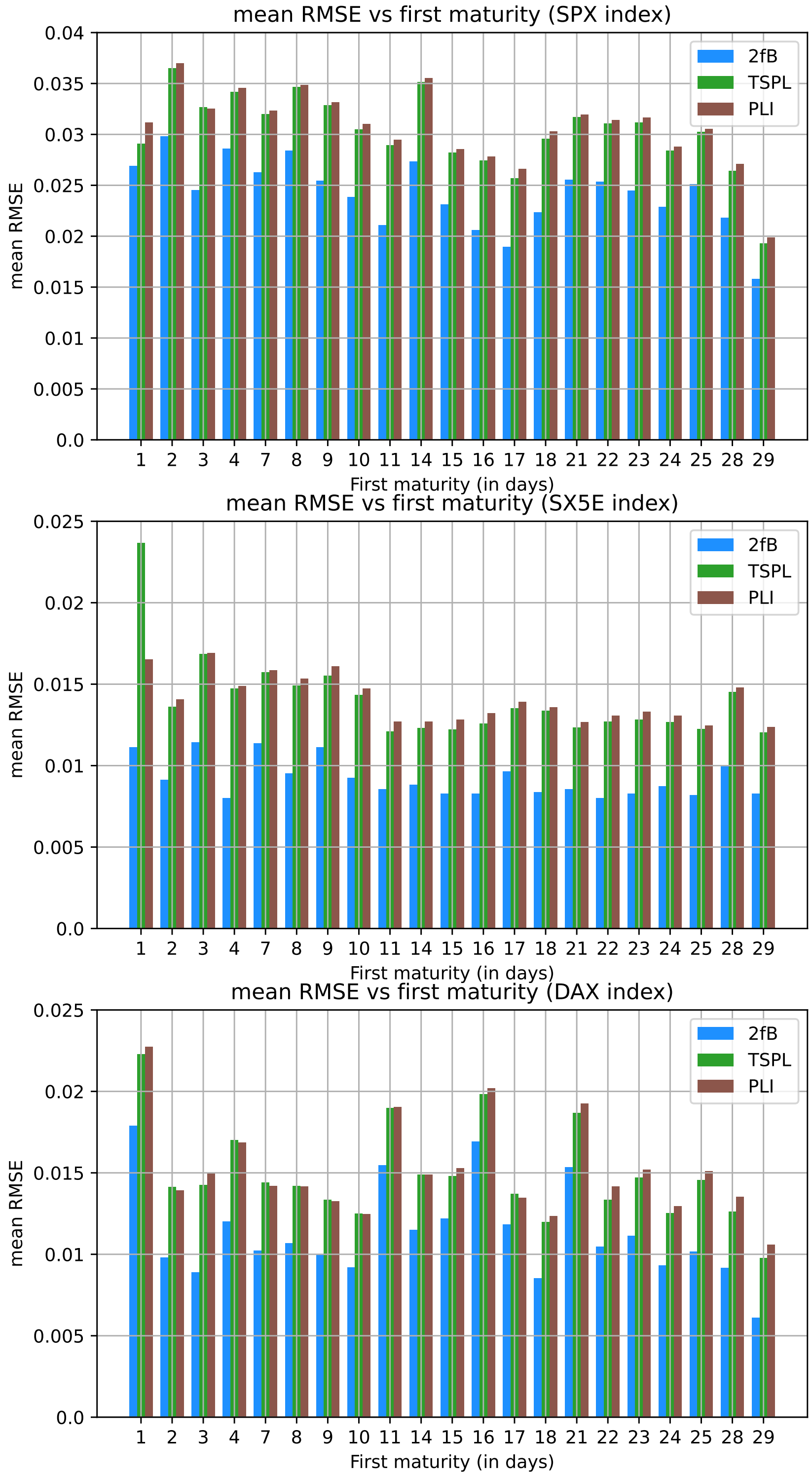

FIGURE 4.6. Mean RMSE as a function of the time to the first monthly maturity (2fB, TSPL and PLI)

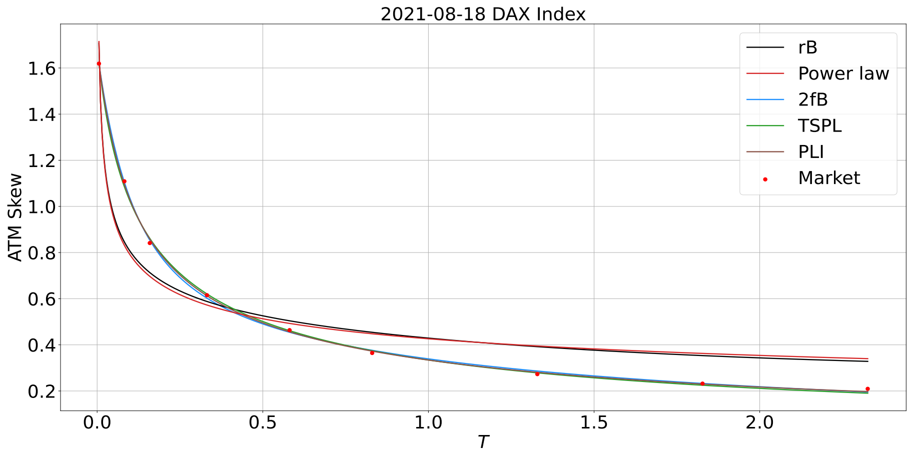

FIGURE 4.7. DAX ATM skew as of August 18, 2021. Comparison of different kernel fits.

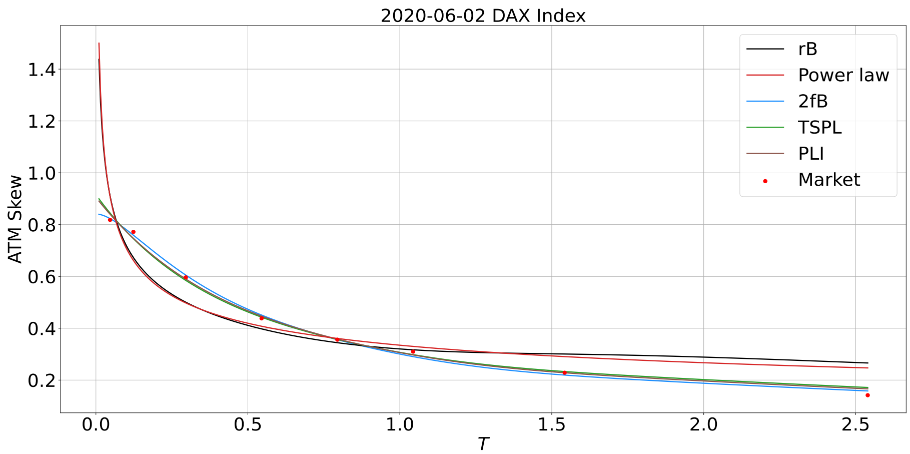

FIGURE 4.8. DAX ATM skew as of June 02, 2020. Comparison of different kernel fits.

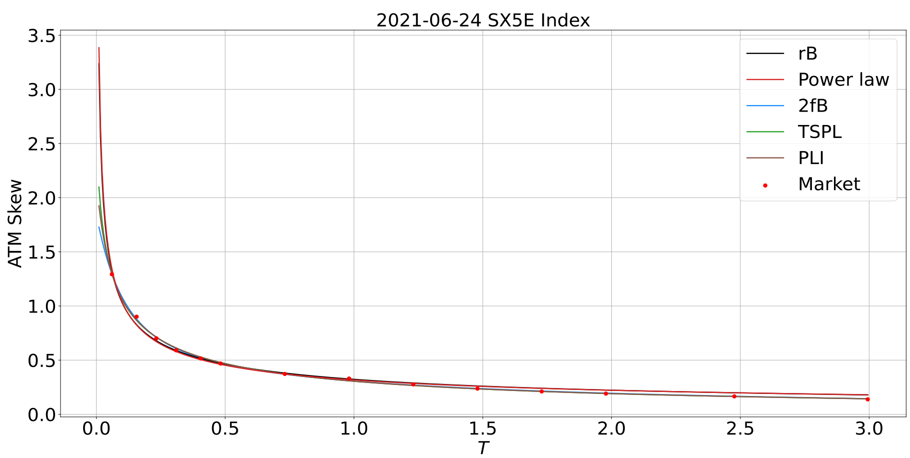

FIGURE 4.9. SX5E ATM skew as of June 24, 2021. Comparison of different kernel fits.

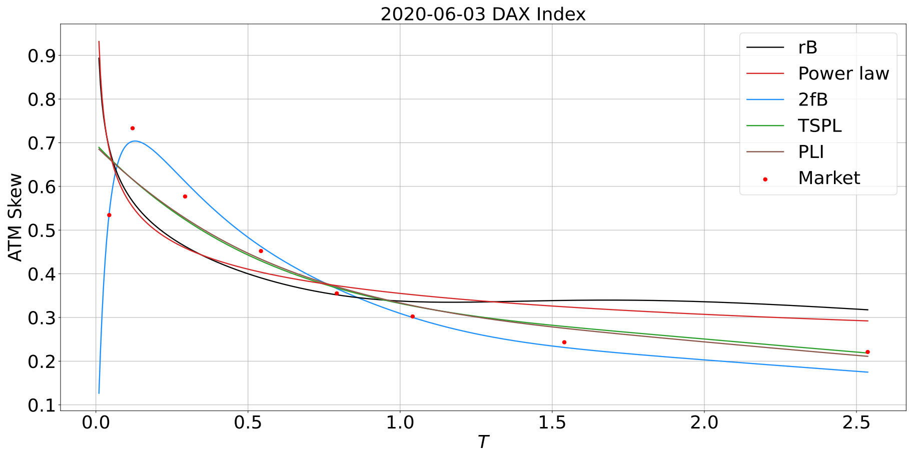

FIGURE 4.10. DAX ATM skew as of June 03, 2020. Comparison of different kernel fits.

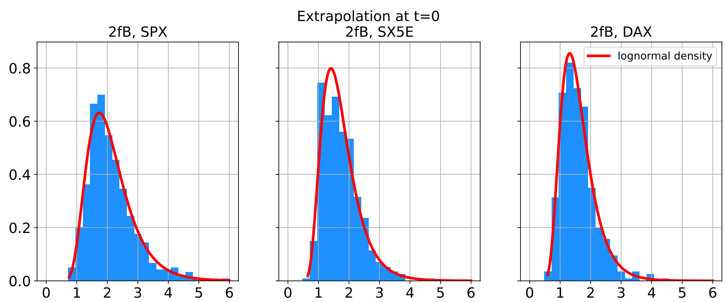

FIGURE 4.11. Extrapolation at $t = 0$ of two-factor Bergomi ATM skew. Red line is a lognormal fit of the density of extrapolation values.

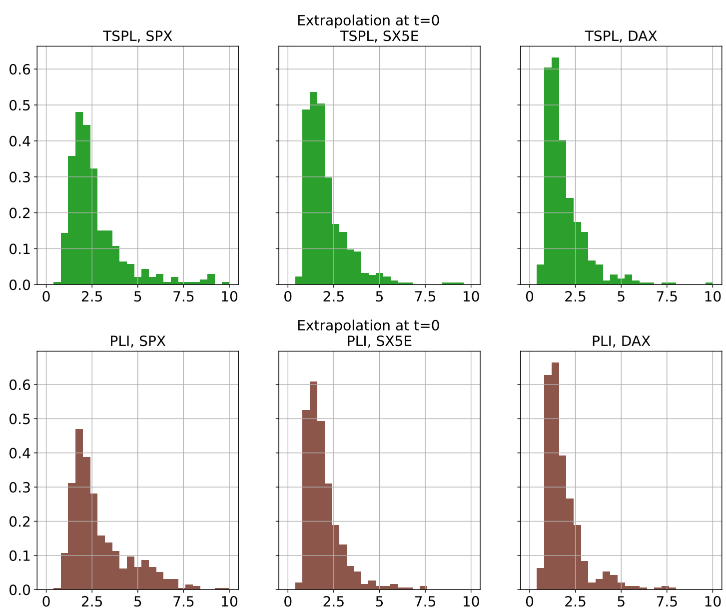

FIGURE 4.12. Extrapolation at $T = 0$ of ATM skew. Top: TSPL. Bottom: PLI

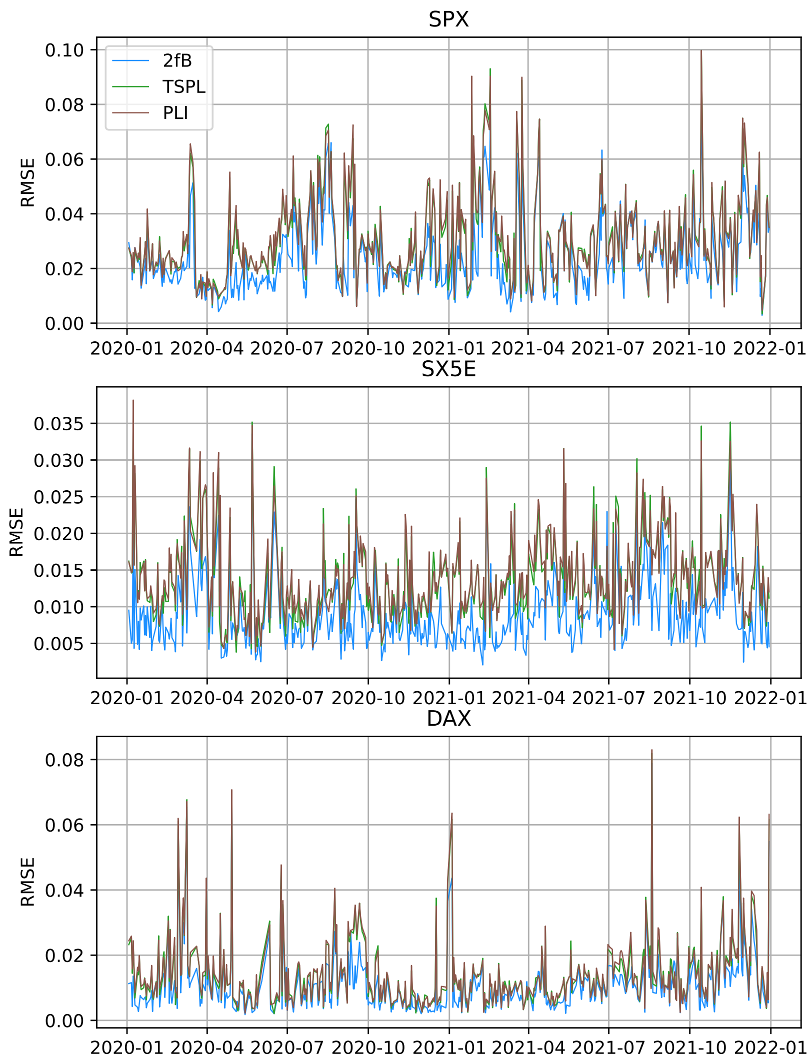

Figure 5.1. Time series of RMSE of 2fB, TSPL and PLI models for each index