Yet another analysis of the SP500 at-the-money skew: crossover of different power-law behaviours

Jules Delemotte, Florent Ségonne, Stefano de Marco

Yet another analysis of the SP500 at-the-money skew: crossover of different power-law behaviours

Jules Delemotte[*1,2], **Stefano De Marco**[1], and **Florent Ségonne**[2]

[1] *CMAP, CNRS, Ecole Polytechnique, Institut Polytechnique de Paris*
[2] *80 Technologies SAS, Paris*

Abstract

Using nine years of options data, we reconstruct a time-series for the at-the-money implied volatility skew $\mathcal{S}_T$ of the SP500 index, and analyse its term structure. Though the well-known power-law dependence $\mathcal{S}_T \approx cT^\alpha$ with $\alpha \approx -0.5$ reported by several works appears to be satisfied for longer maturities, we observe a change of regime in short-dated implied volatilities: the resulting ATM skew behaves, on average, as a superposition of two power laws T^{α_1} for the short-end and T^α for the long-end, with a separation threshold located around two months. The exponent $\alpha_1 \approx -0.3 > \alpha$ points to a smaller explosion rate of the short ATM skew that what would be predicted by longer maturities. We observe a consistent behavior across different parametric and non-parametric estimators, and for both monthly and weekly options. We design a smooth parameterization, that we dub 2 Power Law (having 3 to 5 free parameters), smoothly interpolating between the observed short-term and long-term regimes, which we fit to the daily ATM skew curve on every date in our data set. We observe that the 2 Power Law has performances (in terms of distribution of the mean squared error) equivalent to those of a 2 factor Bergomi parameterization with the same number of parameters. Finally, we study the compatibility of the 2 Power Law model with classical arbitrage bounds on the long term skew, and show how it can be modified for very long maturities so to provide a consistently extrapolation. From a stochastic and notably rough volatility perspective, the observed superposition of power law regimes could motivate models containing two rough factors with different Hurst exponents, or possibly a single rough factor with a varying Hurst exponent. Our independent findings complement the results in the recent work [El Amrani and Guyon, *Does the term structure of Equity at-the-money skew really follow a power law?*, SSRN2002].

Keywords: implied volatility skew, term structure, power-law, arbitrage bounds

1 Introduction

It is a stylized fact reported by several authors [Ber04, Gat06] that the at-the-money (ATM) skew $\mathcal{S}_T = \partial_k \hat{\sigma}(T,k)|_{k=0}$ observed for the implied volatility surface $\hat{\sigma}(T,k)$ of large stock indices such as the SP500 or the Eurostoxx50 is well described by a power-law function of time to maturity T,

$$\mathcal{S}_T^{\mathrm{PL}} = cT^\alpha, \tag{1}$$

where c is a constant, typically negative, and $\alpha \approx -\frac{1}{2}$. Stochastic volatility models used for option pricing and hedging aim at reproducing such a behavior. Classical Markovian models are able to approximately achieve a skew term-structure of the form (1) using at least two factors, one driving the short-end and the other the long-end of the skew curve. For example, the skew term-structure of the well-known two-factor Bergomi model [Ber05], as provided by the Bergomi-Guyon expansion [BG12], is given by a liner

*Corresponding author: jules.delemotte@polytechnique.edu. Date: April 25, 2023

combination of exponential functions of the form $e^{-k\,T}$ (see equation (8) for more details) that has the capability of mimicking a power-law function such as (1) over a wide range of times to maturity, from a few days to several years. More recently, rough volatility models driven by a rough noise with index of Hölder regularity $H \in (0, 1/2)$ have been introduced as (among other features) a parsimonious way to generate a power-law ATM skew, the small-time asymptotics of the skew induced by such models being precisely of the form (1) with $\alpha = H - \frac{1}{2}$, see [ALV07, Fuk11, BFG16]. Among these, the Riemann–Liouville fractional Brownian motion $\int_0^t (t-s)^{H-\frac{1}{2}} dW_s$ driving the variance process in the rough Bergomi model [BFG16] can precisely be seen as a superposition of an infinite number of Markovian factors $\int_0^t e^{-k(t-s)} dW_s$ with different mean reversion speeds k, stretching the idea of the power law being generated by a superposition of factors acting at different time scales.

The divergence of the power-law (1) as $T \to 0$ also opens the question of the explosion (or absence thereof) of the short-term skew, and more generally the question of the possible skew extrapolations below the shortest maturity quoted on the market. The recent work of El Amrani and Guyon [EAG22] precisely addresses this question, by studying which skew model, between the power-law and the 2 factor Bergomi parameterization (or other truncated power-law-like functions) is the more appropriate to describe the ATM skew term structure on a daily basis.

The starting point of our work, which was carried out independently, is close to the one of [EAG22]. Using nine years of options data on the SP500 Index, from February 2007 to the end of 2015, we reconstruct a time-series for the ATM implied volatility skew $\mathcal{S}_T$ using non-parametric (finite difference) and parametric estimators (based on the SVI and SABR models). When analysing the natural logarithm of the average market skew $\log(-\mathbb{E}[\mathcal{S}_T])$ against $\log(T)$, we observe the coexistence of two different behaviors below and above a threshold located around two months: actually, the superposition of two different power laws, since the values of the average skew below and above this change point are approximately located along two lines, with slopes $\alpha_1 \approx -0.3$ for short maturities and $\alpha_2 \approx -0.5$ for longer maturities (this phenomenon has also been reported in a recent update of the preprint [EAG22]). This behavior of the ATM skew curve appears to be consistent across the different skew estimators we have tested (finite difference, SVI, SABR) and across different type of option maturities—monthly or weekly. This observation leads us to postulate the following "2 Power Law model" (2PL)

$$\mathcal{S}_T^{\mathrm{2PL}} = \mathcal{S}_{T_1} \left(\frac{T}{T_1}\right)^{\alpha_1} \left(\frac{\left(\frac{T}{T_1}\right)^{\kappa} + 1}{2}\right)^{\frac{\alpha_2 - \alpha_1}{\kappa}}, \tag{2}$$

which smoothly links the two asymptotic regimes $\mathcal{S}_T^{\mathrm{2PL}} \sim \mathrm{const}\, T^{\alpha_1}$ when $T \to 0$ and $\mathcal{S}_T^{\mathrm{2PL}} \sim \mathrm{const}\, T^{\alpha_2}$ when $T \to \infty$, with a threshold located at $T_1 \approx 2$ months and a bandwith parameter κ. More precisely, our 2PL model (2) is precisely related to the observation that a sigmoid function interpolating between the two asymptotic levels α_1 and α_2 appears to be a good approximation of the log-skew time derivative $\frac{d\log(-\mathbb{E}[\mathcal{S}_T])}{d\log(T)}$, see section 3 for more details.

The 2PL model (2) comes with five parameters (namely $\alpha_1, \alpha_2, \mathcal{S}_{T_1}, T_1, \kappa_1$), but we also formulate restrictions containing four or three free parameters, obtained by anchoring the value of the threshold parameter T_1 and possibly the reference skew $\mathcal{S}_{T_1}$. We perform a daily calibration test of the model (2) to the ATM skew curve for every date in our data set, and observe that the 2 Power Law (together with its 4-parameters version) has performances fully equivalent to those of a 2 factor Bergomi parameterization with the same number of parameters, under different metrics, specifically: the distribution of the calibration error, and the dependence of this error with respect to the first time to maturity.

The 2 Power Law model suggests that the idea of the existence of two components taking place in the skew term-structure and separately driving the short-end and the long-end of the curve, typical of two-factor stochastic volatility models such as the 2-factors Bergomi, could also be cast into the framework of rough volatility models, possibly motivating models containing two rough factors, otherwise a single rough factor with a regularity exponent H varying between two regimes.

Comparison with other works. As pointed out above, a work close in spirit to ours is the aforementioned preprint by Guyon and El Amrani [EAG22]. Our findings are consistent with their results: while [EAG22] advocate a non-explosive parameterization (the 2-factors Bergomi model), we observe that a parameterization

with a tamed explosion rate (the 2 Power Law model) appears to be entirely as appropriate in order to describe the ATM skew curve on a daily basis. In both case, the explosive behavior of the short-end of the skew curve results to be more tempered that what would be predicted by longer maturities. With respect to the analysis carried our in [EAG22], our data set covers a different period (2007–15, containing the subprime crisis, as opposed to 2020-22, containing the Covid-19 crisis), and we also study the impact of weekly options, on top of the most liquid monthly options. More importantly, we introduce the new smooth 2 Power Law model and study its empirical performance, showing it is totally comparable with the one of a 2-factors Bergomi parameterization. As a result, both models could very reasonably be used to extrapolate the ATM skew for very short maturities, and in any case below the first maturity quoted on the market, while leading to different skew asymptotics: finite limiting skew in the 2-factors Bergomi model, and exploding skew in the 2PL model—indicating that the question of the explosive or non-explosive nature of the short-end of the skew curve might be (at least in the case of the SP500 index) hard to disambiguate.

2 Our data set and skew reconstruction

Our data set contains daily values for the last traded call and put option prices and corresponding implied volatilities on the SP500 Index (SPX), from February 2007 to the end of 2015. More precisely, for each trading date we have the values of the last traded implied volatilities on that day (without information on the trade timestamp), for different strikes and several monthly and weekly SPX expirations, with time to maturity up to 900 days. If a certain option is not traded during a certain day, we do not have data for that option on that day. We deduced the value of the forward on the SPX from put-call parity, which we checked for each date and maturity.

For each traded strike and expiry, we are also provided with the total volume (in terms of number of option shares) that has been traded for that particular strike and expiry on each day. We can use these values to inspect option liquidity, providing us with some information about the reliability of the implied volatility data included in our analysis (we refer to the following paragraph and to Appendix B for more information on traded volumes). As another indicator of liquidity, we can look at the number of strikes that are traded for a given expiry at a given date. When the number of traded strikes drops below a certain threshold, we can (and do) consider that the corresponding implied volatility smile is not sufficiently reliable to extract a reasonable value of the ATM skew. With this perspective, we decided to remove from our analysis the smiles with less than five strike points, which correspond to roughly 1.2% of the smiles in our original data set (precisely, 429 smiles out of around 38.000).[1]

Monthly and weekly maturities. Historically, the standard expirations for options on the SPX coincide with the third Friday of each month. Since 2005, the CBOE has also introduced weekly options (or simply *weeklys*), whose expiration dates are not limited to Fridays and to the third week of the month; nowadays, weekly option expire on Monday, Wednesday and Friday of each week.[2] Since their introduction, weekly options have gained popularity and liquidity, in particular for short-dated expirations.

Our data set contains both types of options. While weekly options are rather scarce during the first five years covered by our data, in June 2012 their number increases considerably. This can be seen in Figure 9 in Appendix B, where the top row shows the number of expirations observed on each date: before June 2012 we observe trades for roughly 13 expiries per day (mostly monthly options), whereas the number of observed expirations increases up to 20 or more afterwards (mainly due to the contributions of weeklys). Let us stress that this change of regime in June 2012 is most likely specific to our data set (and probably related to how data have been collected by the provider); it would be difficult to infer a change in the behavior of S&P market participants (precisely located in June 2012) only based on the information contained in Figure 9.

[1] As explained in section 2.1, we first evaluate a reference value for the ATM implied volatility skew using a finite difference estimator, which we then compare to the values obtained via the SVI and SABR parametric models. Setting the threshold to precisely five strikes was mainly dictated by the possibility to fit these parametric models to every smile – the SVI model containing five parameters (and the SABR model four), it did not seem reasonable to perform a fit of SVI to a smile with less than five points.

[2] The symbol for standard monthly options is SPX, whereas for weeklys it is SPXW. More information about these different types of options can be found at https://www.cboe.com/tradable_products/sp_500/spx_options/. The full CBOE option calendar (for 2023) can be found at https://cdn.cboe.com/resources/options/Cboe2023OPTIONSCalendar.pdf.

Monthly options are known to be more liquid than weeklys, and this is also observed in our data set. As an indicator of option liquidity, we computed the product of option prices and traded volumes. More precisely, for each date, we multiply the close price of an option by the number of shares of that option that has been traded on that day. The resulting "dollar volume" variable is shown in Figure 8 for monthly expirations (left) and weekly expirations (right). Using this indicator, monthly options appear to be – on average – ten times more liquid than weekly options (as it can be seen on both the top and middle rows of Figure 8). Nevertheless, weekly options still provide us with additional data to estimate the term structure of the implied volatility skew, notably on the short end, where most of their volume is concentrated (see again Figure 8, middle right plot). If we wish to evaluate the average market ATM skew for a given time-to-maturity over the available observation dates, we can see in Figure 9 (bottom row) that the size of the sample over which the average can be taken is doubled (or more) when taking into account weekly options, for time-to-maturities up to three months.

In order to assess the impact of weekly options on the final results of our analysis, we have evaluated our different skew estimators (introduced in the following sections) using data from monthly options only, from weekly options only, and from both types of options together. Overall, we did not find significant differences between the conclusions that can be drawn in the three cases, as we discuss in the following section.

2.1 Skew reconstruction

For every date t in our data set and for every maturity T observed at the date t, we are provided with values of log-moneyness $k_i = k_i(t,T)$ and implied volatilities $\hat{\sigma}_t(T,k_i)$ for $i \in [1 : n_t^T]$, where n_t^T is the number of observed traded strikes. We aim at estimating the (theoretical) ATM implied volatility skew $\partial_k \hat{\sigma}_t(T,k)|_{k=0}$. It is of course always possible to approximate the theoretical skew with a finite difference estimator of the first order derivative based on the two log-moyeness values closest to the ATM point,

$$\mathcal{S}_T^{\mathrm{FD}}(t) = \frac{\hat{\sigma}_t(T,k_{n_0+1}) - \hat{\sigma}_t(T,k_{n_0})}{k_{n_0+1} - k_{n_0}} \qquad \text{where} \qquad n_0 = \underset{i\in[1:n_t^T]}{\arg\max}\; k_i < 0\,.$$

The quality of the approximation will of course depend on the number n_t^T of available points for each date and maturity. An alternative is to interpolate the values of the implied volatility $\hat{\sigma}_t(T,k_i)$ with a smooth curve $\hat{\sigma}(k)$, and then evaluate the first derivative $\hat{\sigma}'(0)$ exactly on the interpolated curve. We have fitted all the smiles in our data set with two reference models in the industry, the SVI and SABR parameterizations, so to compare the finite difference estimator with the values of the skew yield by these parametric models.

We recall that the SVI model, introduced by Gatheral in [Gat04], is defined by

$$\mathrm{SVI}\left(k;a,b,\rho,\bar{k},\sigma\right) = \left(\hat{\sigma}_t^{\mathrm{SVI}}(T,k)\right)^2 T = a + b\Big(\rho\cdot\left(k-\bar{k}\right) + \sqrt{\left(k-\bar{k}\right)^2+\sigma^2}\Big).$$

The closed-form expression of the ATM skew under the SVI model is $\mathcal{S}_T^{\mathrm{SVI}}(t) = \partial_k \sigma_t^{\mathrm{SVI}}(T,k)\big|_{k=0} = \frac{b\left(\rho - \frac{\bar{k}}{\sqrt{\bar{k}^2+\sigma^2}}\right)}{2\left(a-b\left(\rho\bar{k}-\sqrt{\bar{k}^2+\sigma^2}\right)\right)}$. We fit the parameters a, b, ρ, $\bar{k}$ and σ for each date t and maturity T using a weighted Least Squares (LS) objective

$$a_t^T, b_t^T, \rho_t^T, \bar{k}_t^T, \sigma_t^T = \underset{a,b,\rho,\bar{k},\sigma\in D}{\arg\min} \sum_{i=1}^{n_t^T} \Big(\mathrm{SVI}\left(k_i;a,b,\rho,\bar{k},\sigma\right) - \left(\hat{\sigma}_t(T,k_i)\right)^2 T\Big)^2 w(k_i), \tag{3}$$

where $D = \{0 \le a \le T\max_i\left(\hat{\sigma}_t(T,k_i)\right)^2, b > 0, \rho \in (-1,1), \min\left(2\min_i k_i, 0\right) \le \bar{k} \le \max\left(2\max k_i, 0\right), \sigma > 0\}$ and $w(k) \propto \exp\left(-\frac{k^2}{\delta}\right)$ are Gaussian weights with kernel bandwidth of the form $\delta = c\sqrt{T-t}$, so to increase the importance of the implied volatility values around the ATM point.[3] We solve the LS problem (3) applying the quasi-explicit minimization algorithm described in [DMM09]. The SVI model is known to have good calibration performances, and indeed we find a very good fit for essentially all the smiles in our data

[3] We look for $k_t(\tau) = \inf\{k_i < 0 : \frac{\sigma_t(t+\tau,0)}{\sigma_t(t+\tau,k_i)} \ge \frac{1}{2}\}$. A regression against $\log(\tau)$ shows that the average curve $\tau \mapsto |k(\tau)|$ is well approximated by the function $c\sqrt{T-t}$ with $c = 0.038$.

set. In order too deal with the no-butterfly arbitrage condition (not included in the definition of the set D and in the quasi-explicit algorithm of [DMM09]), we have applied the no-arbitrage check described in [MM22] ex-post to all the fitted SVI smiles.[4] We only keep the smiles free of butterfly arbitrage, corresponding to 97% of the total smiles in the data set (roughly 36000 smiles), and remove the others from our analysis.

For the SABR model, introduced by Hagan et al. [HKLW02], we use the formulation of the smile parameterization in [Obl08], which we recall in Appendix A for the reader's convenience. We fit the four parameters α, β, ν, ρ of the smile formula (17) by minimization of the same weighted LS objective (3) as for the SVI model. It is well known that the SABR formula can be affected by butterfly arbitrage, and the identification of the admissible parameter set is even more difficult that for SVI, due to the more involved expression of the SABR smile (17). Violations of the no-butterfly arbitrage condition have been reported in particular for low strikes (corresponding to log-moneyness below -0.25 in [HKLW14]); this drawback of the SABR formula should arguably not be a major issue for our purpose of ATM skew reconstruction. We evaluate the ATM skew using centered finite differences on the (deterministic) smile formula (17), see (18). Though we also obtain a good fit of most smiles with the SABR model, we observe that, not surprisingly, the 4-parameters SABR formula has on average smaller accuracy than the 5-parameters SVI model, in particular for shorter maturities (below one month).

Having estimated a value of the ATM skew for every observation date and every maturity accessible at that date, we are provided with a sample of observations of the market ATM skew for times to maturity ranging between 1 day and 2.5 years. The sample size highly depends on time to maturity. For times to maturity below three months, the sample of observed ATM skews contains up to one hundred points, while being limited to ten observations for times to maturity between 2 and 2.5 years (see Figure 9 for more details), when including monthly options. When including weekly options, the number of observations for times to maturity below three months considerably increases, up to two hundreds (see again Figure 9, bottom left panel). As as start, we can study the behavior of the average (or median) skew as a function of time to maturity, while postponing to the next section an analysis of the *daily* behavior of the ATM skew curve.

Remark 1. In what follows, we still denote T the time to maturity, and $\mathbb{E}[\mathcal{S}_T]$ the market skew averaged over the observation dates.

The results obtained with the finite difference (FD) skew estimator applied to monthly options are shown in Figure 1. In the top plot, we observe the typical power-law-like behavior of the negative average market skew $-\mathbb{E}[\mathcal{S}_T]$. The bottom plot displays the natural logarithm of the average skew $\log(-\mathbb{E}[\mathcal{S}_T])$ against $\log(T)$. As expected from the power-law-like behavior, the dependency is approximately linear, but this second plot also reveals a finer structure in the behavior of the ATM skew below and above a threshold located around two months: the values of the average log skew below and above this change point are approximately distributed along two lines, with two different slopes. This phenomenon (also reported in a recent update of the preprint [EAG22]) is more apparent in Figure 2, where the slopes of the linear regressions performed separately for times to maturity below and above two months are provided.

Different skew estimators. In Appendix C.1, we show the results obtained when estimating the ATM skew with the SVI and SABR models as described above, and compare them with the finite difference estimator. Figure 10 (resp. Figure 11) shows the average (resp. the median) of the opposite skew $-\mathcal{S}_T$ against T, on log-log plots, using the samples of ATM skews obtained with the three different estimators. One can see that the results are consistent, in the sense that all three methods provide a clear distinction between the short and long end of the ATM skew curve, and the slopes of linear regressions for long-term maturities are the same (up to two digits) in the three cases. Not surprisingly, the largest differences are observed in the short-term, where we can observe a bias between the average (or median) skews obtained with the different methods. For maturities below one month, the SVI and SABR models are calibrated to smiles with a small number of strikes, which results in a considerable degree of uncertainty in the determination of the ATM skew. For the SVI model, the overall short-term slope (-0.32) eventually turns out to be consistent with the one for the FD estimator (-0.31). As pointed out above, in our tests we observe that the SABR formula has lower calibration accuracy than the SVI model, in particular for short-dated smiles, which causes a sizeable difference between the short-term slope of the average SABR skew (-0.37) with respect to the

[4] We might have applied the no butterfly arbitrage condition of [MM22] within the fitting procedure, so to output an arbitrage-free SVI parameter set for every market smile, but this would have been too time-consuming, since the calibration of a single arbitrage-free SVI smile using the procedure described in [MM22] takes around 30 seconds in our tests.

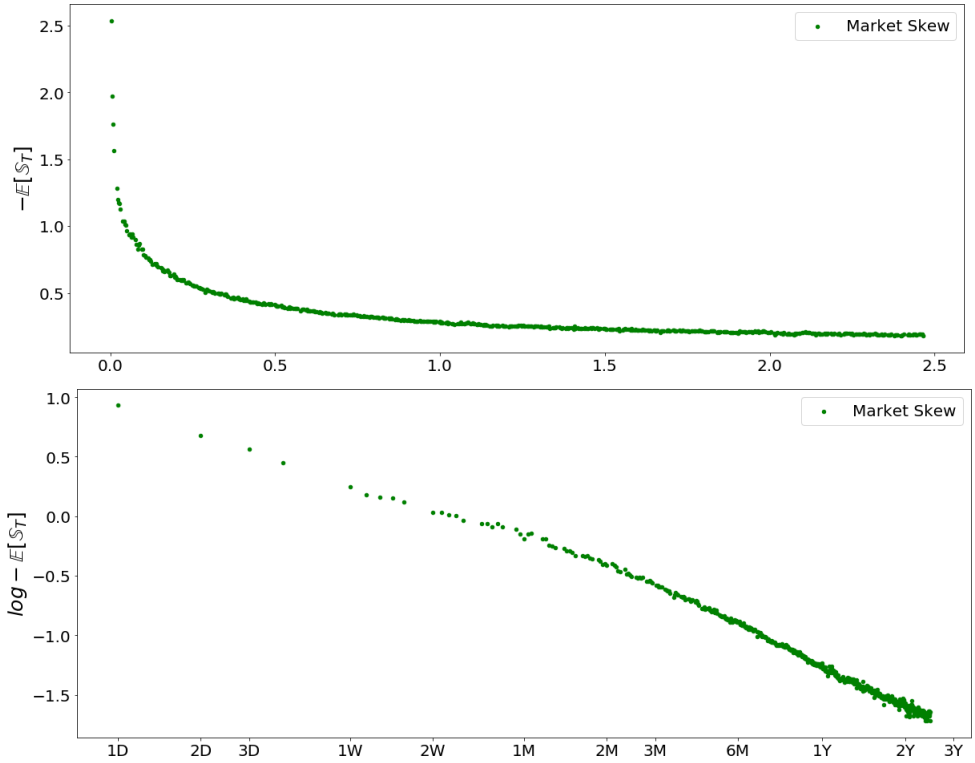

Figure 1: **Term structure of the average ATM skew**. *Top figure*: opposite of the market skew reconstructed by finite difference, averaged over all observation dates in our data set, against time to maturity. *Bottom figure*: same as above, on a log-log plot. Only options with monthly maturities are taken into account (in this figure; see Appendix C.2 for additional plots including weekly maturities).

other estimators. For completeness, we also provide the slopes that the short-term linear regressions have when we do not take into account the point corresponding to the shortest time to maturity (equal to 1 day), for which the largest amount of uncertainty is observed: the resulting values are -0.31 (FD), -0.29 (SVI), and -0.35 (SABR).

Impact of weekly options. Figure 1 only takes into account the data from options with standard monthly maturities. In Figure 12 in Appendix C.2, we show the results obtained when injecting weekly options in addition to monthly options (rightmost plot), or when considering weekly options only (middle plot). We observe an interesting persistence of the phenomenon of superposition of two linear behaviors with different slopes for the ATM skew, with a threshold always located around two months, independently of the type of options taken into account.

3 Parametric models for the ATM skew term-structure

A power-law dependence of the average skew with respect to time to maturity,

$$-\mathbb{E}\left[\mathcal{S}_T\right]^{\mathrm{PL}} = cT^{\alpha}, \tag{4}$$

for some $c > 0$ and $\alpha < 0$ would of course result in a straight line with slope α in the bottom plot of Figure 1. A best least-squares fit of the parameterization (4) to the data in Figure 1 yields $\alpha = -0.44$. This value is mainly driven by long-term maturities, the number of points with time to maturity above two months being considerably larger than the number of points below this threshold.

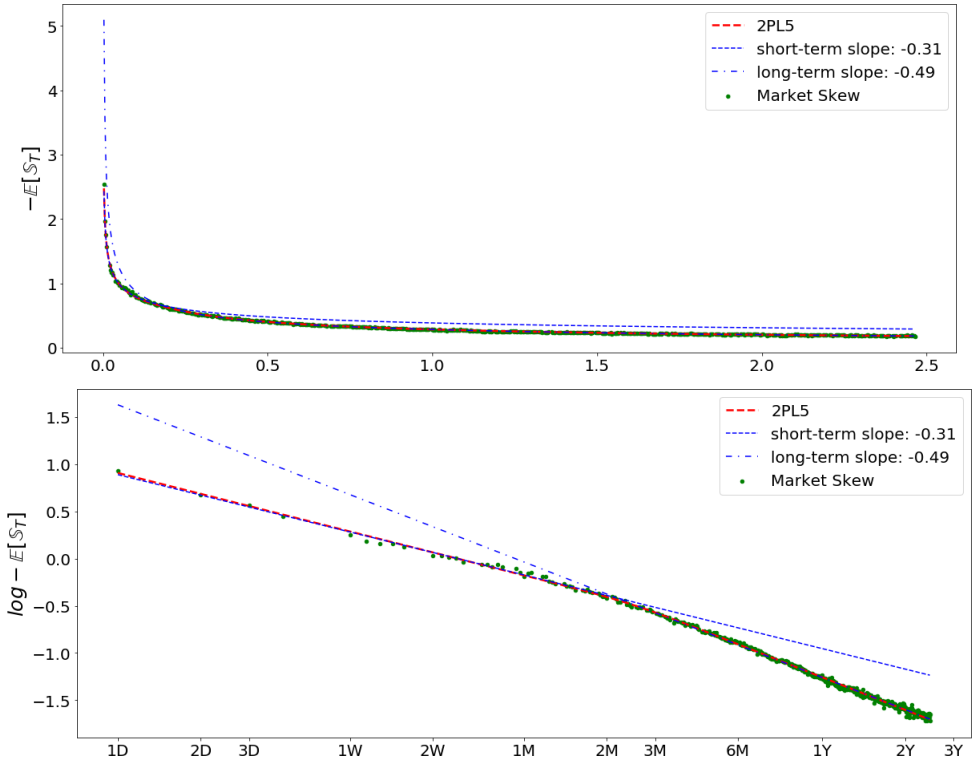

Figure 2: **Average ATM skew, short-term and long-term linear behaviors and the smooth 2 Power Law (2PL) model**. As in Figure 1, the green dots represent observations of the market ATM skews from monthly options, averaged over all dates in our data set, plotted against time to maturity on a normal scale (top plot) and on a log-log scale (bottom plot). The blue dashed or dash-dotted lines represent linear regressions performed separately for maturities below and above 2 months. The red dashed line is a best fit of our smooth 2 Power-Law model (7) with five parameters (2PL5). See section 3 for other instances of the 2PL model containing fewer parameters.

As pointed out above, Figure 1 suggests that a behavior of the form

$$-\mathbb{E}\left[\mathcal{S}_T\right] \propto \begin{cases} T^{\alpha_1} & \text{for } T < \frac{60}{365} \\ T^{\alpha_2} & \text{for } T > \frac{60}{365} \end{cases} \qquad \text{with } \alpha_2 \leq \alpha_1 < 0$$

would describe more closely the observed average market skew. We could therefore postulate a parameterization of the form

$$\log\big(-\mathbb{E}\left[\mathcal{S}_T\right]\big) = (\beta_1 + \alpha_1 \log(T))\mathbb{1}_{T<T_1} + (\beta_2 + \alpha_2 \log(T))\mathbb{1}_{T>T_1}, \tag{5}$$

where T_1 is a threshold between the short-term and long-term linear behaviors. The parameterization (5) is not a smooth function of T—we could of course impose conditions on the parameters α_i, β_i and T_1 in such a way that the right hand side of (5) is continuous at T_1, but it would not be smooth (unless $\alpha_1 = \alpha_2$). We find it more natural to design a smooth parameterization, with unconstrained parameters.

The smooth 2 Power-Law model (2PL). In Appendix D, we show a local estimate of the log-skew derivative $\frac{d\log(-\mathbb{E}[\mathcal{S}_T])}{d\log(T)}$, obtained from linear regressions of the values of $\log(-\mathbb{E}\left[\mathcal{S}_T\right])$ over a moving window. We find that a sigmoid function smoothly interpolating between the asymptotic values α_1 and α_2 fits the log skew derivative reasonably well (as it can be see in Figure 13), motivating a parameterization (for the time derivative) of the form

$$\frac{d\log(-\mathbb{E}\left[\mathcal{S}_T\right])}{d\log T} = \alpha_1 + (\alpha_2 - \alpha_1)\,\sigma\big(\kappa(\log T - \log T_1)\big), \tag{6}$$

	$\mathcal{S}_{T_1}$	T_1	κ	α_1	α_2
2PL	-0.639	0.185	5.03	-0.310	-0.494

Table 1: Parameters of the 2PL skew model (7) fitted to the average ATM skew term-structure in Figure 1

	c_1	k_1	c_2	k_2
2fB	-1.212	2.94	-4.177	211

Table 2: Parameters of the 2-factors Bergomi skew model (8) fitted to the average ATM skew term-structure in Figure 1

where

$$\sigma(x) = \frac{e^x}{1+e^x}, \qquad x \in \mathbb{R},$$

and $\kappa > 0$ is a constant tuning the bandwidth of the smooth transition induced by the sigmoid function. When κ becomes large, the right hand side of (6) approaches a piece-wise constant function with a jump at $T = T_1$. We can integrate the right-hand side of (6) between T_1 and T: interestingly, taking advantage of the properties of the sigmoid function, we obtain a closed-form expression for the average skew, given by

$$\mathbb{E}\left[\mathcal{S}_T\right]^{\text{2PL}} = \mathcal{S}_{T_1}\left(\frac{T}{T_1}\right)^{\alpha_1}\left(\frac{\left(\frac{T}{T_1}\right)^{\kappa}+1}{2}\right)^{\frac{\alpha_2-\alpha_1}{\kappa}} := \mathcal{S}_T^{\text{2PL}}, \tag{7}$$

which we dub 2 Power Law (2PL) parameterization, where $\mathcal{S}_{T_1}$ represents the (negative) value of the skew at time to maturity T_1. It is easy to see that $\mathcal{S}_T^{\text{2PL}} \sim \text{const}\, T^{\alpha_1}$ when $T \to 0$ and $\mathcal{S}_T^{\text{2PL}} \sim \text{const}\, T^{\alpha_2}$ when $T \to \infty$. A least square fit (with uniform weights) of the 2PL (7) to the average skew term-structure is shown in Figure 2, where we can see that the fit is excellent. The calibrated parameters are given in Table 1, where the value of T_1 corresponds to 67 days.

Reducing the number of parameters (from the 2PL5 *to the* 2PL4 *and the* 2PL3*).* The 2 Power Law model (7) contains five parameters, which of course makes it much more versatile than the single power-law (4) (which is actually embedded in (7) as a special case). In order to make a fair comparison with other skew parameterizations based on a smaller number of parameters (such as the one induced by the two factors Bergomi model, recalled in the next paragraph), we can consider some restrictions of the 2PL functional form $\mathcal{S}_T^{\text{2PL}}$. We can, for example, fix once and for all the value of the threshold to $T_1 = \frac{60}{365}$: the resulting parameterization depends on the remaining four parameters $(\alpha_1, \alpha_2, \mathcal{S}_{T_1}, \kappa)$ and will therefore be labeled 2PL4 in our tests. We can further restrict the model by anchoring the graph of the 2PL function to a specific point: fixing the value of $\mathcal{S}_{T_1}$, we can anchor the 2PL4 to the point $(T_1, \mathcal{S}_{T_1})$, obtaining a family of functions depending on the remaining three parameters $(\alpha_1, \alpha_2, \kappa)$. In practice, when fitting the 2PL to the market skew term-structure observed on a given day (as we do in the following section), we can replace $(T_1, \mathcal{S}_{T_1})$ in (7) with the point $(T^{\text{mkt}}, \mathcal{S}_{T^{\text{mkt}}})$, where T^{mkt} is the first observed time to maturity above $\frac{60}{365}$ and $\mathcal{S}_{T^{\text{mkt}}}$ the observed ATM skew for that maturity. We label the resulting three-parameters model 2PL3. This procedure can for example be followed when traders want to have zero error for a specific maturity when fitting the market skew term-structure. To stress the difference with its restricted versions, from now on the most general five-parameters version of the 2PL model (7) will be labeled 2PL5.

2-factors Bergomi model (2fB). The two-factors forward variance model of Bergomi [Ber08] admits an explicit approximation for the ATM implied volatility skew, achieved via the so-called Bergomi-Guyon expansion [BG12]. The resulting ATM skew approximation has the following form:

$$\mathcal{S}_T^{\text{2fB}} = c_1 \frac{k_1 T - \left(1 - e^{-k_1 T}\right)}{\left(k_1 T\right)^2} + c_2 \frac{k_2 T - \left(1 - e^{-k_2 T}\right)}{\left(k_2 T\right)^2}, \tag{8}$$

depending on the four parameters c_1, c_2 and $k_1, k_2 > 0$. Even if the asymptotic behavior of the function on the right hand side of (8) for large and small T is different from the one of (7)—it is easy to see that

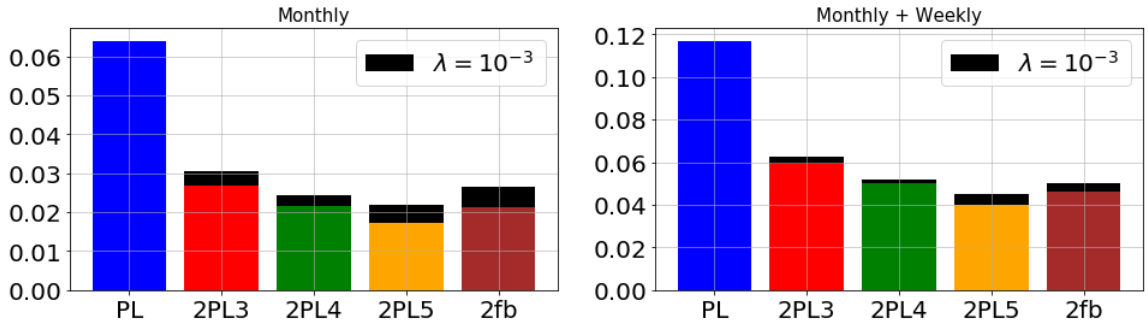

Figure 3: **Average daily error (RMSE) for the different skew parameterizations**. Average RMSE for different models fitted to the ATM skew curve observed on every date in our data set (after 01-01-2012, when weekly options are more abundant). *Left plot*: monthly options only (roughly ten observed maturities per day). *Right plot*: monthly and weekly options (roughly twenty observed maturities per day). The different models are: the single Power-Law PL (4), the 2 Power-Law (7) with different numbers of parameters (2PL5, 2PL4 and 2PL3), and the 2-factors Bergomi (2fB). In each column, the coloured part represents the average RMSE for a fit of the model without penalization (corresponding to $\lambda = 0$), while the black upper part represents the additional calibration error introduced by the regularized objective in (9) with $\lambda = 10^{-3}$.

$\mathcal{S}_T^{2\text{fB}} \sim$ const when $T \to 0$ and $\mathcal{S}_T^{2\text{fB}} \sim$ const T^{-1} when $T \to \infty$—the 2fB parameterization has the remarkable property of being able to mimick the behavior of the power law function $T^{-\alpha}$ with $\alpha \approx -0.5$ over a wide range of times to maturity T, from a few days to a few years, by choosing two very distinct values of the characteristic times $1/k_1$ and $1/k_2$ (as seen in Table 2). A fit of the 2fB model (8) to the average ATM skew term-structure in Figure 1 is accurate (definitely more precise than the single power law (4)), but still less accurate than the 2PL (7) shown in Figure 2, due to the lack of the appropriate asymptotically linear behaviors for $\log(-\mathbb{E}\left[\mathcal{S}_T\right]^{2\text{fB}})$ for large and small T.

3.1 Fitting accuracy on a daily basis

We now fit the power-law model (4), the 2PL5 model (7) and the 2fB model (8) on a daily basis, that is, to the ATM skew curve observed on every single pricing date contained in our dataset. As mentioned above, we also consider the restricted version 2PL4 with four parameters (where we fix $T_1 = \frac{60}{365}$) and the 2PL3 with three parameters (where we fix $(T_1, \mathcal{S}_{T_1}) = (T^{\text{mkt}}, \mathcal{S}_{T^{\text{mkt}}})$, T^{mkt} being the first observed time to maturity above $\frac{60}{365}$ and $\mathcal{S}_{T^{\text{mkt}}}$ the observed ATM skew). On each date, the calibration is performed by minimization of a LS objective with uniform weights. We should keep in mind that the number of maturities observed on a given date is much smaller than the number of points contained in Figure 2 (which takes into account all dates in the data set). As seen in Figure 9 (top plots), at a given date we observe on average around 10 maturities for monthly options and around 20 maturities when taking into account weekly options as well. When fitting a model with four or five parameters to ten points on a daily basis, we can of course expect some instabilities in the calibration output.[5] In order to increase the stability and interpretability of the calibrated parameters, we penalize the least square objective as follows

$$\theta_t = \underset{\theta \in D}{\arg\min} \frac{1}{n_t} \sum_{i=1}^{n_t} \left(\mathcal{S}_{T_{t,i}^{\text{mkt}}} - f(\theta, T_{t,i}^{\text{mkt}}) \right)^2 + \frac{\lambda}{dim(D)} \left\| \frac{\theta - \overline{\theta}}{\overline{\theta}} \right\|^2 \tag{9}$$

where θ_t represents the vector of model parameters ($\theta_t = (\alpha_1, \alpha_2, T_1, \mathcal{S}_{T_1}, \kappa)$ for the 2PL5, $\theta_t = (c_1, c_2, k_1, k_2)$ for the 2fB, and so on), $f(\theta, T)$ the parametric model for the skew curve, and $T_{t,i}^{\text{mkt}}$ are the n_t maturities

[5]Even if the 2fB and 2PL models have better performances than a single power-law or a one factor Bergomi model, on some dates, we encouter situations where the former models appear to be over-parameterized, since on those specific dates the observed ATM skew curve is actually well described by a single power-law, or by only one of the two terms of the sum in (8). In these situations, some of the model parameters saturate to their limiting values—$k_1 = 0$ or $k_2 \to \infty$ for the 2fB, $T_1 = 0$ for the 2PL—producing spikes in the time series of calibrated parameters.

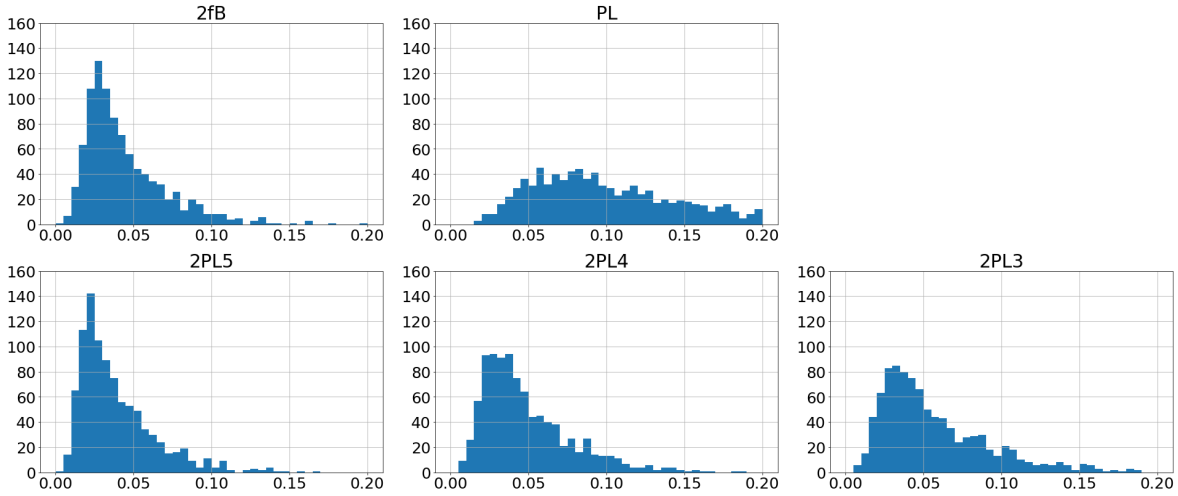

Figure 4: **Distribution of the daily RMSE**. Empirical distribution of the RMSE resulting from the calibration to the daily ATM skew term-structure of monthly and weekly options. PL: Power-Law, 2fB: 2-factors Bergomi model, 2PLn: instances of the 2 Power-Law model with n free parameters.

observed at time t. In the penalization term, we penalize the relative difference with respect to the vector $\overline{\theta}$ of model parameters calibrated to the average ATM skew (reported in Tables 1 and 2 for the 2PL and 2fB models). The normalization according to the number $dim(D)$ of parameters is used in order to compare models having different number of parameters while keeping the same regularization parameter λ.

Figure 3 shows the resulting average Root Mean Squared Error (RMSE) of the different models. Some comments are in order:

1. Of course, we expect a model with five our four (or three) parameters to perform better than the single power-law (4) with two parameters, which is what we see in Figure 3. Our objective is to observe the relative performances of the different parameterizations. From this perspective, the 2fB and 2PL models perform on average two to three times better than the single power-law PL: as also illustrated in [EAG22], the power-law function seems to be too poor to appropriately capture all the information contained in the market skew term structure.

2. At the same time, another message conveyed by Figure 3 is that, overall, the PL model does *not* perform so bad: it would still seem reasonable to use the single power-law parameterization (4) as a first approximation of the daily ATM skew curve. The more advanced 2PL model can be used when further accuracy is desired, in particular for shorter maturities.

3. Reducing the number of free parameters of the 2PL model only slightly deteriorates the fitting accuracy. The 2PL4 model has performance essentially equivalent to the 2fB model having the same number of parameters (and the 2PL3 remains comparable with both). Looking at Figure 3 (or at the histograms of the RMSE in Figure 4), it would be difficult to conclude which model, between the 2fB and the 2PL, is more appropriate to describe the ATM skew curve on a daily basis.
   If we are concerned with the question of the extrapolation of the skew for very short maturities, and in any case below the first maturity quoted on the market, the 2fB and the 2PL models lead to different skew asymptotics: finite limiting skew $\frac{1}{2}(c_1 + c_2)$ in the 2fB model, and exploding skew in the 2PL model.

4. The calibration error for monthly and weekly options together (left plot in Figure 3) is larger than the error for monthly options alone (right plot), because the number of maturities targeted by the calibration to monthly and weekly options is larger (roughly twice as large).

In addition to the average RMSE, in Figure 4 we show the empirical distribution of the daily RMSE of the different models. As expected, the distribution of the RMSE of the power-law is more spread out than

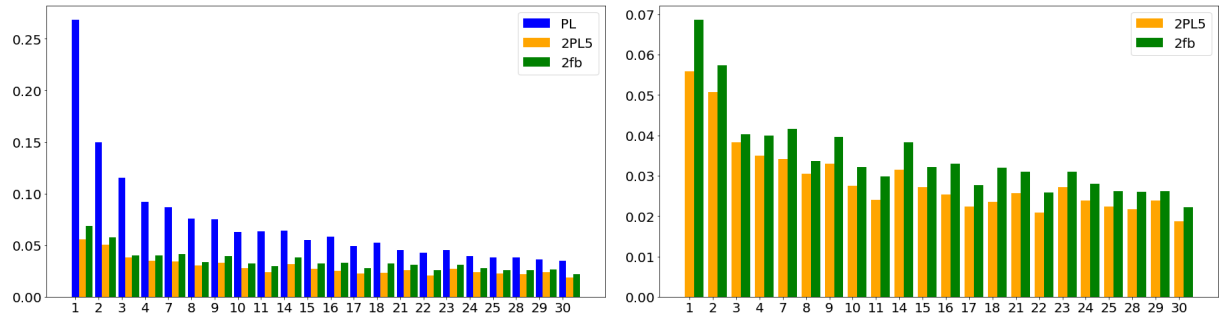

Figure 5: **Dependence of the RMSE with respect to first time to maturity**. Mimicking an idea in [EAG22], we show the average RMSE for the PL, 2PL5 and 2fB models models fitted to the daily ATM skew, displayed according to the value of the first time to maturity, ranging from 1 to 30 days.

the one of the 2PL and the 2fB model. The error distributions for the 2fB model and the 2PLn models are very similar.

Time series of calibrated parameters. The time series of the parameters of the 2PL5 and the 2fB models calibrated via the regularized objective (9) for dates between 2012 and 2015 are shown in Appendix E. Overall, even if the daily time series are rather volatile (as we can expect from the calibration of a four or five parameters model to a curve containing roughly twenty points), most of the parameters remain within a well-identified and interpretable range. For the 2fB model, it is well known that the parameters $k_1 \approx 3$ and $k_2 \approx 200$ are linked to two different characteristic times $\tau_1 = \frac{1}{k_1} \approx 4$ months and $\tau_2 = \frac{1}{k_2} \approx 2$ days, respectively related to the long and short-end of the skew curve.[6] For the 2PL5 model in Figure 15, we can see that the average values taken by the long and short power-law exponents $\alpha_2 \approx -0.5$ and $\alpha_1 \approx -0.25$ are in line with what expected from Figure 2. We observe that the threshold parameter T_1 takes values in a quite tight range (between $0.04 \approx 14$ days and $0.25 \approx 90$ days). The parameter $\mathcal{S}_{T_1}$ represents the value of the skew at time to maturity T_1 and therefore follows the realized movements of the market skew. As already mentioned above, the 2PL and the 2fB models are both related to the idea of two components acting in the skew term-structure. Figures 15 and 16 point to the existence of these two components in the daily ATM skew: one driving short maturities (represented by the parameter α_1 for the 2PL model, resp. k_1 for the 2fB model), and the other driving longer maturities (represented by the parameter α_2, resp. k_2). The 2PL5 model also materialises the transition between the two regions, via the threshold parameter T_1.

Dependence of the RMSE with respect to first time to maturity. Mimicking an idea in [EAG22], in Figure 5 we show the average calibration error of the PL, 2PL5 and 2fB models as a function of the first time to maturity, the first maturity being a standard monthly.[7] The results in Figure 5 are fully consistent with [EAG22]: the calibration error of the PL model has a significant dependence with respect to the first time to maturity—pointing again to the fact that a single power-law appears to be appropriate to describe the ATM skew for maturities above one month (where the errors of the three models reach a comparable size), but struggles to describe both the short- and long-end of the curve. On the contrary, both the 2PL and 2fB models have the capability of fitting the ATM skew in the short and long ends simultaneously, their level of accuracy being only weakly dependent on first time to maturity. Consistently with Figures 3 and 5, we can see from the right plot of Figure 5 that the 2fB and the 2PL models essentially display the same calibration performances (with a slight advantage for the 2PL, also due to the fact that we have considered the 2PL5 instance of the model here).

Remark 2. The 2PL parameterization of the skew can be plugged into a model for the reconstruction of the volatility surface. The popular SSVI model of Gatheral and Jacquier's [GJ14] is parameterized by the ATM

[6] We see from Figure 16 that sometimes k_2 take very high values, close to 800, associated to the particularly short maturity $\frac{365}{800} \approx 0.4$ days. One could set a maximum value for k_2 equal to 365: we also performed the calibration test of the 2fB model imposing this restriction on k_2, and this did not affect considerably the accuracy of the output.

[7] More precisely, in Figure 5, we take into account the ATM skews of monthly and weekly options, in the following way: for each pricing date, we remove the weekly options before the first observed monthly maturity, and then consider all observed skews (monthly and weeklys) after the first monthly option. Doing so, we have a first time to maturity that ranges from 1 to 30 days, for otherwise, using all available weeklys options the first time to maturity would always be below 10 days.

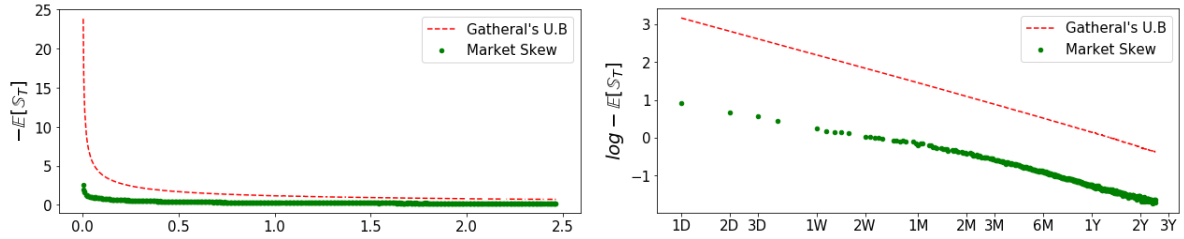

Figure 6: **Average ATM skew and no-arbitrage upper bound**. Green dots: ATM skew from monthly options, averaged over all dates in our data set, against time to maurity (same data as in Figure 1). Red dashed line: the upper-bound (10) computed using the average ATM implied volatility over the same dates. Right plot: same as left plot, on a log-log scale.

total variance curve $\theta_T = \sigma(T,0)^2 T$, by a function $\varphi : \mathbb{R}_+ \to \mathbb{R}_+$ and a parameter ρ. The ATM implied volatility skew under this model is given by

$$\mathcal{S}_T^{\text{SSVI}} = \frac{\rho\sqrt{\theta_T}}{2\sqrt{T}}\varphi\left(\theta_T\right).$$

Setting the 2PL parameterization (7) as the target skew term-structure allows to drive the choice for the function φ. Under the simple form $\theta_T = \sigma^2\, T$ of the total variance curve, a function φ precisely given by the 2PL function, that is $\varphi(\theta) = \text{const} \times \left(\frac{\theta}{\theta_1}\right)^{\alpha_1}\left(\frac{\left(\frac{\theta}{\theta_1}\right)^{\kappa}+1}{2}\right)^{\frac{\alpha_2-\alpha_1}{\kappa}}$, leads to the desired behavior for the implied volatility skew. See [GJ14, Corollary 4.1 and Remark 4.4] for hints on the arbitrage-freeness of an SSVI model with such a structure.

4 Arbitrage Bounds on the long-term skew

No-arbitrage bounds on the at-the-money implied volatility skew were presented in Hodges [Hod96], and then refined by Gatheral in [Gat00]. The lower bound on the ATM skew (that is, an upper bound for the opposite of the skew) in [Gat00] reads

$$-\mathcal{S}_T(t) \leq \frac{1}{\sqrt{T}}\frac{N\left(-\frac{1}{2}\hat{\sigma}_t(T,0)\sqrt{T}\right)}{N'\left(-\frac{1}{2}\hat{\sigma}_t(T,0)\sqrt{T}\right)} := \text{UB}(\hat{\sigma}_t(T,0),T), \tag{10}$$

where $N(x) = \int_{-\infty}^{x} e^{-y^2/2}\frac{dy}{\sqrt{2\pi}}$ denotes the standard Gaussian cdf. Since the total volatility $\hat{\sigma}_t(T,k)\sqrt{T}$ vanishes as time to maturity T tends to zero, it is easy to see that, on the one hand, we have $\text{UB}(\hat{\sigma}_t(T,0),T) \underset{T\to 0}{\sim} \sqrt{\frac{\pi}{2T}}$. On the other hand, the asymptotic behavior of the upper bound in the long-term depends on the behavior of the total volatility for large maturities. Assuming that the ATM implied volatility $\hat{\sigma}_t(T,0)$ remains bounded away from zero as $T \to \infty$ (which is the case in particular if $\hat{\sigma}_t(T,0)$ has a non-zero limit $\hat{\sigma}_t^{\infty}$), then the total implied volatility diverges, and

$$\text{UB}(\hat{\sigma}_t(T,0),T) \underset{T\to\infty}{\sim} \frac{2}{\hat{\sigma}_t(T,0)T} \tag{11}$$

follows from the well-known asymptotics $\frac{N(z)}{N'(z)} \sim \frac{1}{|z|}$ as $z \to -\infty$.

It is clear that the asymptotic upper bound above is not consistent with any parameterization implying a power-law behavior of the skew of the form $T^{-\frac{1}{2}}$, or more generally T^{α} for some $\alpha \in (-1,0)$. In particular, the 2PL parameterization (7) will violate the upper bound UB beyond some time to maturity T^*, since $\frac{\mathcal{S}_T^{\text{2fB}}}{\text{UB}(\hat{\sigma}_t(T,0),T)} \underset{T\to\infty}{\sim} \text{const}\,\hat{\sigma}_t(T,0)T^{1+\alpha_2}$, and the latter function tends to infinity as T becomes large, for any value of $\alpha_2 \in (-1,0)$ (in particular then for $\alpha_2 \approx -\frac{1}{2}$).

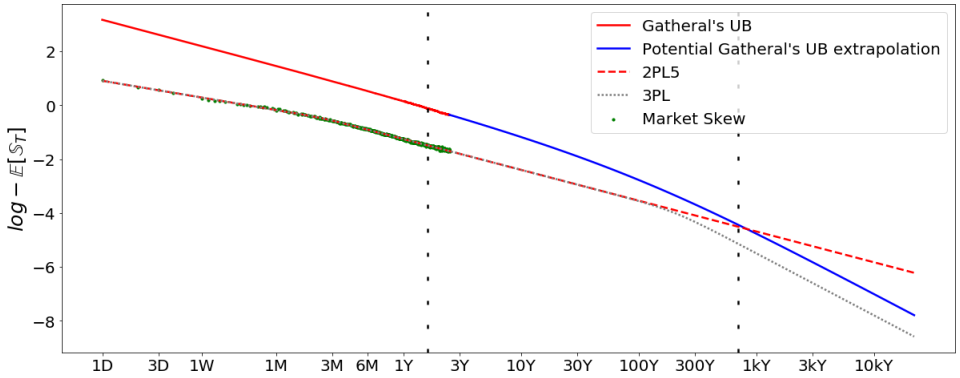

Figure 7: **Skew extrapolation for long maturities and the consistent 3PL model.** We extend the 2PL5 parameterization (red dashed line) and the modified 3PL model (14) (grey dashed line) calibrated to the average ATM market skew (green dots) beyond the range of traded market maturities, together with the skew upper bound (10) (blue line) under a flat ATM implied volatility extrapolation.

Of course, for every date t one can evaluate the actual upper bound (10) by injecting the value of the ATM implied volatility $\hat{\sigma}_t(T,0)$ observed on that day into the expression of UB (the market ATM skew $-\mathcal{S}_T^{\text{mkt}}(t)$ is trivially expected to satisfy (10), if the data come from an implied volatility surface free of static arbitrage). In Figure 6, we compare the average market ATM skew (the same as in Figure 1) with the function obtained injecting the average ATM implied volatility $\frac{1}{n}\sum_{i=1}^{n}\hat{\sigma}_{t_i}(T,0)$ into the upper bound (10). We can see that the market skew lies well below the upper bound, over the range of times to maturity observed in our data set. In the log-log plot (right pane of Figure 6), the upper bound appears to behave almost linearly with a slope close to $-\frac{1}{2}$, which shows that we are still far from the large-time asymptotics (11) (and actually rather close to the short-time asymptotic regime of UB) for expiries up to three years .

In Figure 7, we extend the upper bound (10) beyond the largest observed time to maturity $T_{\max} \approx 2.5\,y$ using a flat extrapolation of the ATM implied volatility $\hat{\sigma}(T,0)$ (that is $\hat{\sigma}(T,0) = \hat{\sigma}(T_{\max},0)$ for $T \geq T_{\max}$). We can observe that, as expected, the 2PL parameterization (7) eventually crosses the upper bound, but this only happens for very large maturities, around 700 years. Roughly in the same region (for times to maturity over one thousand years), the upper bound is seen to reach its asymptotic regime (11) of the form $\frac{c}{T}$.

4.1 Consistent extrapolation of the long-term skew: modification of the 2PL parameterization

Figure 14 shows the result of adding an additional sigmoid function $\sigma\big(\kappa_2(\log T - \log T_2)\big)$ to the log skew derivative with respect to the parameterization introduced in (6):

$$\frac{d\log(-\mathbb{E}\left[\mathcal{S}_T\right])}{d\log T} = \alpha_1 + (\alpha_2-\alpha_1)\,\sigma\big(\kappa_1(\log T - \log T_1)\big) + (\alpha_3-\alpha_2)\,\sigma\big(\kappa_2(\log T - \log T_2)\big), \tag{12}$$

where we have set $\kappa_1 = \kappa$ and we impose the constraint $\alpha_3 \leq -1$. The second sigmoid function term in (12) essentially does not perturb the values taken by the first one, while introducing a smooth transition to the asymptotic value α_3 for large T, provided that T_2 is sufficiently far away from T_1—in practice, it is enough to have

$$T_1 e^{5/\kappa_1} < T_2 e^{-5/\kappa_2}. \tag{13}$$

As in the case of the 2PL model (7), equation (12) can also be integrated exactly. We obtain a 3 Power Law (3PL) model asymptotically consistent with the no-arbitrage upper bound (10):

$$\mathbb{E}\left[\mathcal{S}_T\right]^{\text{3PL}} = \mathcal{S}_{T_1}\left(\frac{T}{T_1}\right)^{\alpha_1}\left(\frac{\left(\frac{T}{T_1}\right)^{\kappa_1}+1}{2}\right)^{\frac{\alpha_2-\alpha_1}{\kappa_1}}\left(\frac{\left(\frac{T}{T_2}\right)^{\kappa_2}+1}{\left(\frac{T_1}{T_2}\right)^{\kappa_2}+1}\right)^{\frac{\alpha_3-\alpha_2}{\kappa_2}} := \mathcal{S}_T^{\text{3PL}} \tag{14}$$

or yet

$$\mathcal{S}_T^{\mathrm{3PL}} = \mathcal{S}_T^{\mathrm{2PL}} \left(\frac{\left(\frac{T}{T_2}\right)^{\kappa_2} + 1}{\left(\frac{T_1}{T_2}\right)^{\kappa_2} + 1} \right)^{\frac{\alpha_3 - \alpha_2}{\kappa_2}},$$

which shows that the 3PL model only adds a multiplicative term to the initial 2PL function.

Figure 7 shows (in the grey dashed curve) the 3PL function (14) with parameters $\alpha_1, \alpha_2, \mathcal{S}_{T_1}, T_1, \kappa_1$ identical to the 2PL parameterization (see Table 1), and additional parameters $\alpha_3 = -1$, $T_2 = 200$ and $\kappa_2 = 3.3$. As expected, while being superposed to the 2PL model over the range of traded market maturities, for times to maturity around T_2 the 3PL model bends over so to reach an asymptotically linear behavior of the form $cT^{-\alpha_3} = cT^{-1}$. Note that, while the parameters shared with the 2PL (namely $\theta_1 = (\alpha_1, \alpha_2, \mathcal{S}_{T_1}, T_1, \kappa_1)$) are intended to be calibrated to market data, the additional parameters $(\alpha_3, T_2, \kappa_2)$ can be freely chosen by the user—in such a way to respect condition (13). Our reference value for the (very) long-term asymptotic slope is $\alpha_3 = -1$, but other values of $\alpha_3 \leq -1$ could be specified. Of course, the restrictions of θ_1 to the four-parameters or three-parameters models 2PL4 and 2PL3 presented in section 3 can still be considered.

Remark 3 (Generalization to the nPL). Iterating the procedure described above, the 2PL skew model (7) could be further extended to a nPL model

$$\mathbb{E}\left[\mathcal{S}_T\right]^{\mathrm{nPL}} = C\, T^{\alpha_1} \prod_{i=1}^{n-1} \left(\left(\frac{T}{T_i}\right)^{\kappa_i} + 1 \right)^{\frac{\alpha_{i+1} - \alpha_i}{\kappa_i}}, \tag{15}$$

with pillar maturities $(T_0 = 0, T_1, \ldots, T_{n-1}, T_n = \infty)$ and approximate power-law behaviors with exponent α_i on each of the intervals (T_{i-1}, T_i). The functional form (15) contains $3n - 1$ parameters in total: $(T_i, \kappa_i, \alpha_i)_{1 \leq i \leq n-1}$, together with α_n and the constant C. In case one wants to specify the value of skew $\mathcal{S}_{T_j}$ for a reference maturity T_j as one of the model parameters, it is possible to use the following equivalent formulation:

$$\mathbb{E}\left[\mathcal{S}_T\right]_j^{\mathrm{nPL}} = \mathcal{S}_{T_j} \left(\frac{T}{T_j}\right)^{\alpha_1} \prod_{i=1}^{n-1} \left(\frac{\left(\frac{T}{T_i}\right)^{\kappa_i} + 1}{\left(\frac{T_j}{T_i}\right)^{\kappa_i} + 1} \right)^{\frac{\alpha_{i+1} - \alpha_i}{\kappa_i}}. \tag{16}$$

The formulation (7) for the 2PL model (resp. (14) for the 3PL model) is recovered setting $n = 2$ (resp. $n = 3$) and $j = 1$. Of course, while being particularly versatile, the nPL model (15) leads to an arguably indesirable proliferation of parameters, alterating the original simplicity of a skew parameterization based on one or two power laws.

Acknowledgements. JD would like to thank the quantitative research team at 80 Technologies for all the advice and technical help with data processing. SDM thanks J. Gatheral for stimulating insights, and gratefully acknowledges financial support from the research projects *Chaire Risques Financiers* (École Polytechnique, Fondation du Risque and Société Générale) and *Chaire Stress Test, Risk Management and Financial Steering* (École Polytechnique, Fondation de l'École Polytechnique and BNP Paribas).

References

[ALV07] E. Alos, J.A. Leon, and J. Vives. On the short-time behavior of the implied volatility for jump-diffusion models with stochastic volatility. *Finance and Stochastics*, 11:571–589, 2007.

[Ber04] L. Bergomi. Smile Dynamics. *Risk*, pages 117–123, September 2004.

[Ber05] L. Bergomi. Smile Dynamics ii. *Risk*, pages 67–73, October 2005.

[Ber08] L. Bergomi. Smile Dynamics III. *Risk*, pages 90–96, October 2008.

[BFG16] C. Bayer, P. Friz, and J. Gatheral. Pricing under Rough Volatility. *Quantitative Finance*, 16(6):887–904, 2016.

[BG12] L. Bergomi and J. Guyon. Stochastic volatility's orderly smiles. *Risk*, pages 60–66, October 2012.

[DMM09] S. De Marco and C. Martini. Quasi-Explicit Calibration of Gatheral's SVI model. Zeliade Systems WhitePaper, 2009.

[EAG22] M. El Amrani and J. Guyon. Does the term structure of equity at-the-money skew really follow a power law ? *https://ssrn.com/abstract=4174538*, 2022.

[Fuk11] M. Fukasawa. Asymptotic analysis for stochastic volatility: Martingale expansion. *Finance and Stochastics*, 15:635–654, 2011.

[Gat00] J. Gatheral. Rational Shape of the Volatility Surface. *Conference at Quant Congress USA, Boston*, 2000.

[Gat04] J. Gatheral. A parsimonious arbitrage-free implied volatility parameterization with application to the valuation of volatility derivatives. *Presentation at Global Derivatives & Risk Management, Madrid*, 2004.

[Gat06] Jim Gatheral. *The volatility surface: a practitioner's guide*. John Wiley & Sons, 2006.

[GJ14] J. Gatheral and A. Jacquier. Arbitrage-free SVI volatility surfaces. *Quantitative Finance*, 14(1):59–71, 2014.

[HKLW02] P.S. Hagan, D. Kumar, A.S. Lesniewski, and D.E. Woodward. Managing Smile Risk. *Wilmott magazine*, 1:84–108, 2002.

[HKLW14] P.S. Hagan, D. Kumar, A.S. Lesniewski, and D.E. Woodward. Arbitrage-free SABR. *Wilmott magazine*, 1:60–75, 2014.

[Hod96] H.M. Hodges. Arbitrage Bounds on the Imlied Volatility Strike and Term Structures of European-Style Options. *The Journal of Derivatives*, 3(4):23–35, 1996.

[MM22] C. Martini and A. Mingone. No arbitrage svi. *SIAM Journal on Financial Mathematics*, 13(1):227–261, 2022.

[Obl08] J. Obloj. Fine-tune your smile: Correction to Hagan et al. `https://arxiv.org/abs/0708.0998`, 2008.

A SABR smile formula

We recall the SABR formula for the implied volatility surface:

$$\sigma(T,k) = I\left(\alpha,\beta,\nu,\rho,K,F,T\right) = I_0\left(k\right)|_{k=\log\left(\frac{K}{F}\right)}\left(1 + T\, I_1(y)|_{y=KF}\right), \tag{17}$$

where

$$I_0\left(k=\log\left(\frac{K}{F}\right)\right) = \begin{cases} \alpha K^{\beta-1} & \text{if } K = F \text{ i.e. } k=0 \\ -\nu \dfrac{k}{\log\left[\frac{z-\rho+\sqrt{1+z^2-2z\rho}}{1-\rho}\right]} & \text{if } K \neq F \text{ i.e. } k \neq 0, \end{cases}$$

$z(K,\beta,\alpha,\nu,F) = \nu\frac{F^{1-\beta}-K^{1-\beta}}{\alpha(1-\beta)}$, and

$$I_1(y) = \frac{1}{24}y^{\beta-1}\alpha^2(\beta-1)^2 + \frac{1}{4}y^{\frac{\beta-1}{2}}\alpha\beta\nu\rho + \frac{1}{24}\nu^2\left(2-3\rho^2\right).$$

For each date t and maturity T, we evaluate the ATM SABR skew using centered finite differences

$$\mathcal{S}_T^{SABR}(t) = \frac{I\left(\alpha,\beta,\nu,\rho,F+10^{-1},F,T\right) - I\left(\alpha,\beta,\nu,\rho,F-10^{-1},F,T\right)}{2\cdot 10^{-1}}F\,. \tag{18}$$

B Options liquidity

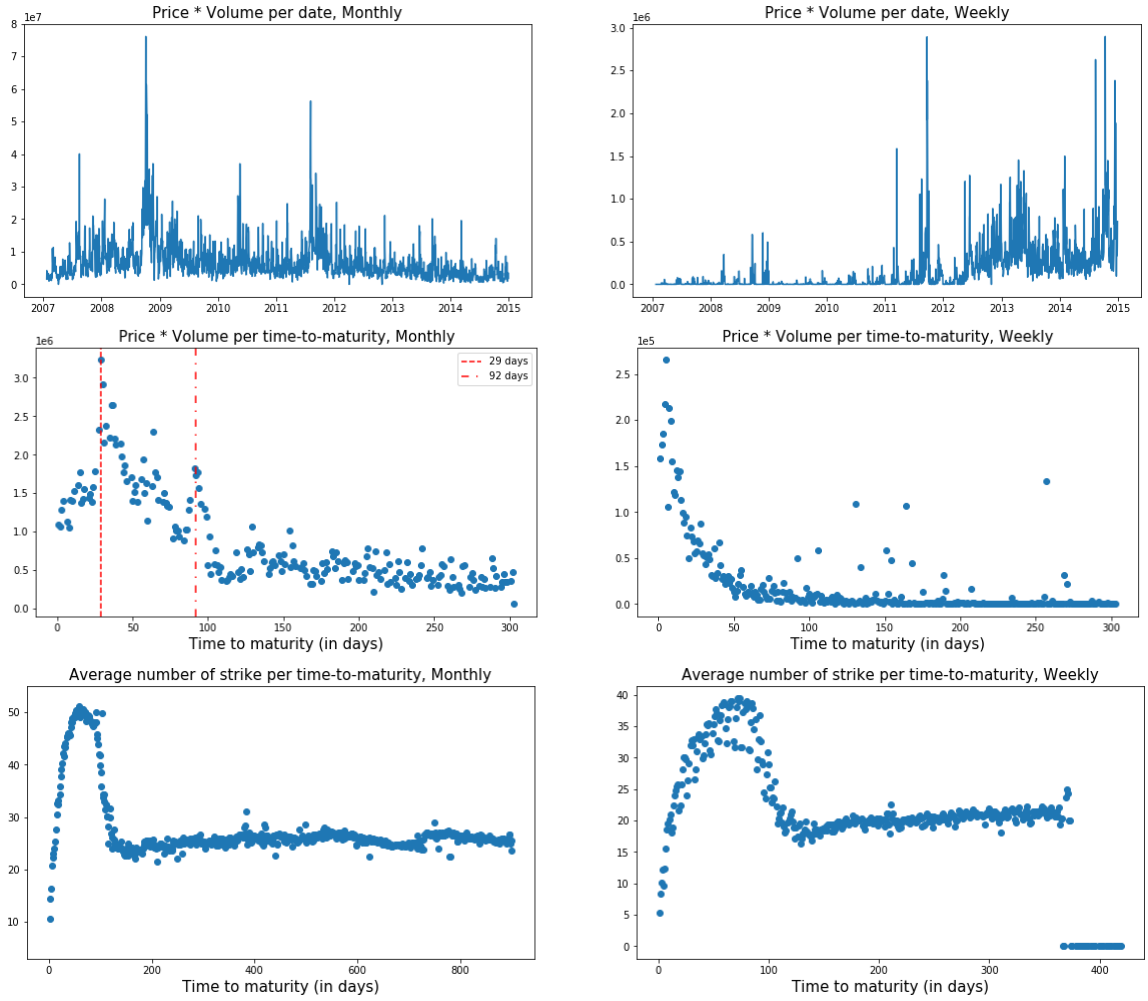

Figure 8: **Traded volumes and traded strikes.** *Top row*: we show the time series of the "dollar volume" variable (option price multiplied by traded volume), totaled over all the options traded on each day. *Middle row*: we plot the term-structure of the same variable, that is the total "dollar volume" for each time-to-maturity (totaled over traded strikes), averaged over the observation dates. *Bottom row*: as another indicator of liquidity, we plot the number of traded strike as a function of time-to-maturity, averaged over the observation dates. In all cases, we separate monthly options (left) and weekly options (right).

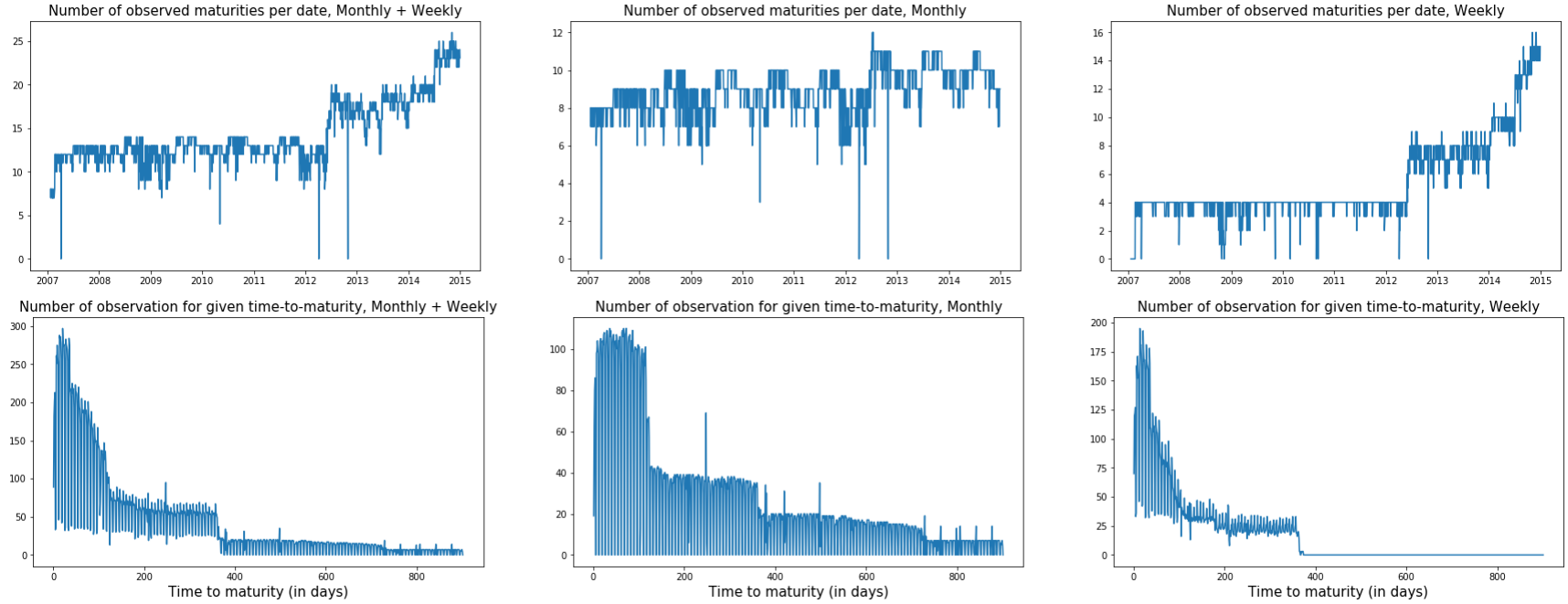

Figure 9: **Traded maturities**. *Top row*: for each date in our data set, we plot the number of different expirations for which a trade is seen. *Bottom row*: for every n between 1 and 900, we plot the number of observation dates where a time-to-maturity of n days appears in the data set. In the middle and right columns, we distinguish between the different types of expiry (monthly or weekly).

C Comparison of different skew reconstruction methods

C.1 Different skew estimators

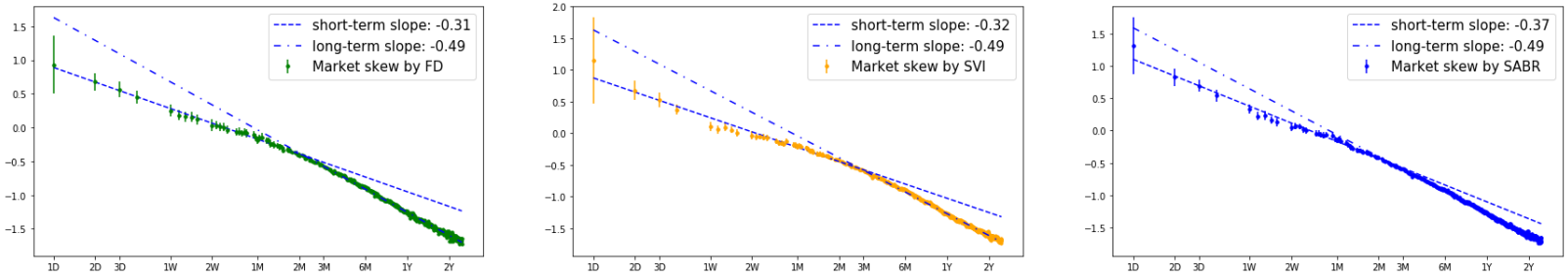

Figure 10: **Average ATM skew with parametric and non parametric estimators**. As in Figure 1, we plot the empirical average of the opposite skew $-\mathbb{E}[\mathcal{S}_T]$ against time to maturity T, on a log-log plot. From left to right, the daily skew $\mathcal{S}_T$ is evaluated by finite difference, with the SVI formula and with the SABR formula calibrated to market implied volatilities as explained in section 2.1. In all cases, the error bars display the 95% confidence interval for the empirical mean (they are only shown for maturities below two months; the error bars for larger maturities are very small). When removing the shortest maturity equal to 1 day, the slopes of the short-term linear regressions become (from left to right): -0.31 (FD), -0.29 (SVI) and -0.35 (SABR).

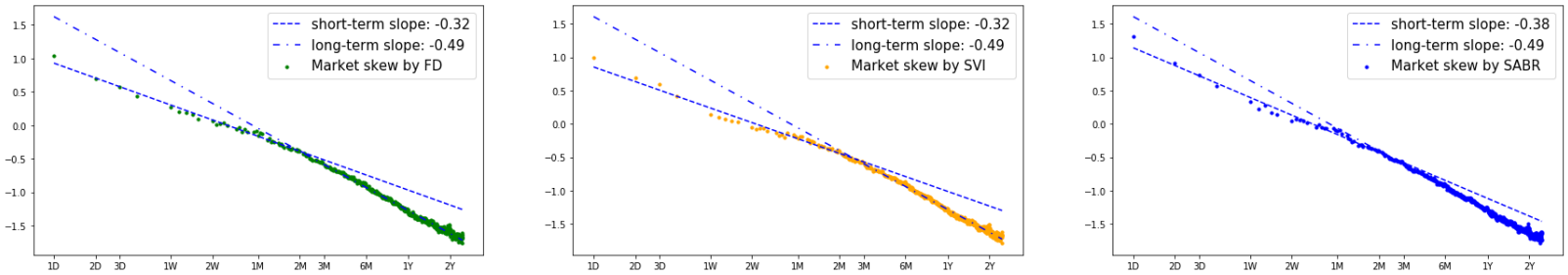

Figure 11: **Median ATM skew with different estimators**. As a complement to Figure 10, we plot the median (as opposed to the average) of the observed ATM skews $-\mathcal{S}_T$ against time to maturity T (again on a log-log plot), using the same samples of values of $\mathcal{S}_T$ used in Figure 10 (evaluated, from left to right, by finite difference, with the SVI model, and with the SABR model).

C.2 Impact of weekly options

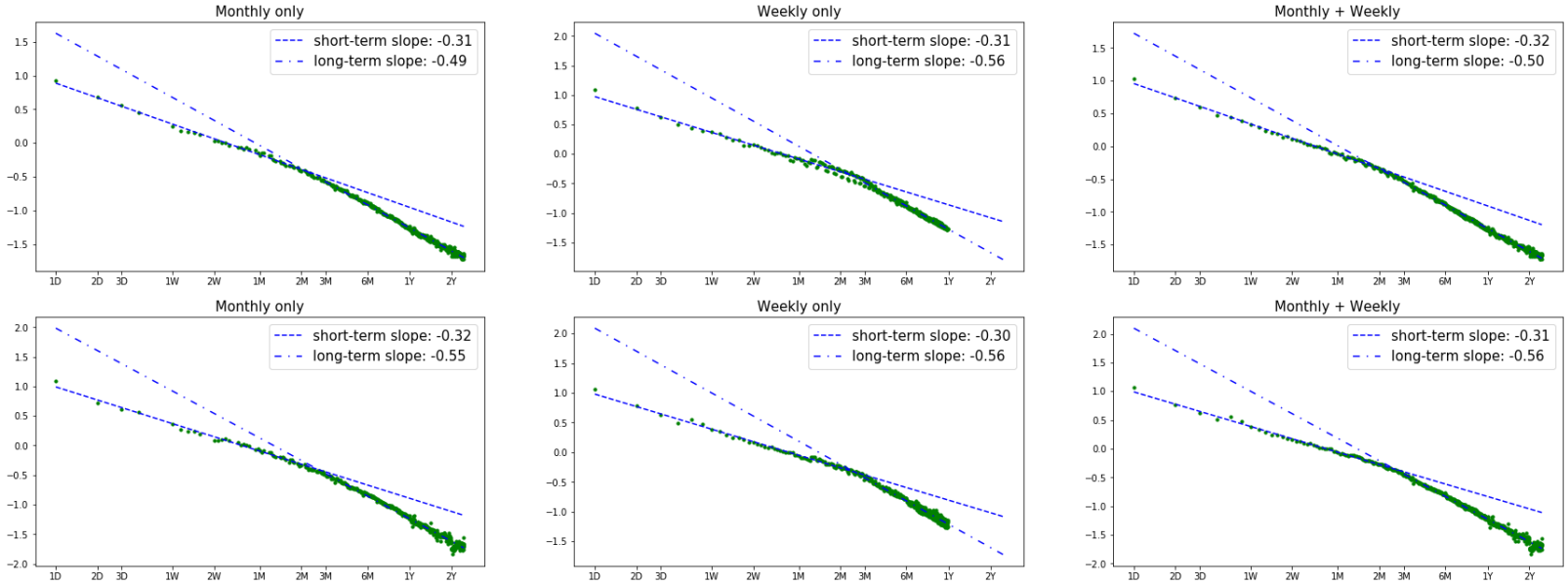

Figure 12: **Monthly and weekly options**. As in Figure 1, we plot the empirical average of the opposite skew $-\mathbb{E}[\mathbb{S}_T]$ against time to maturity T, on a log-log plot, where the skew $\mathbb{S}_T$ is evaluated with a finite difference estimator. *Top plots*: the average is taken over all dates in our data set, from to 2007 to 2015. From left to right, we take into account only monthly options (the plot is therefore identical to 1), only weekly options, and both types of options together. *Bottom plots*: the average is taken over observation dates between 2012 and 2015, the period where weekly options are most abundant in our data set, as seen in Figure 9, top right plot.

D Log skew time derivative and relation with the 2 (or 3) Power Law

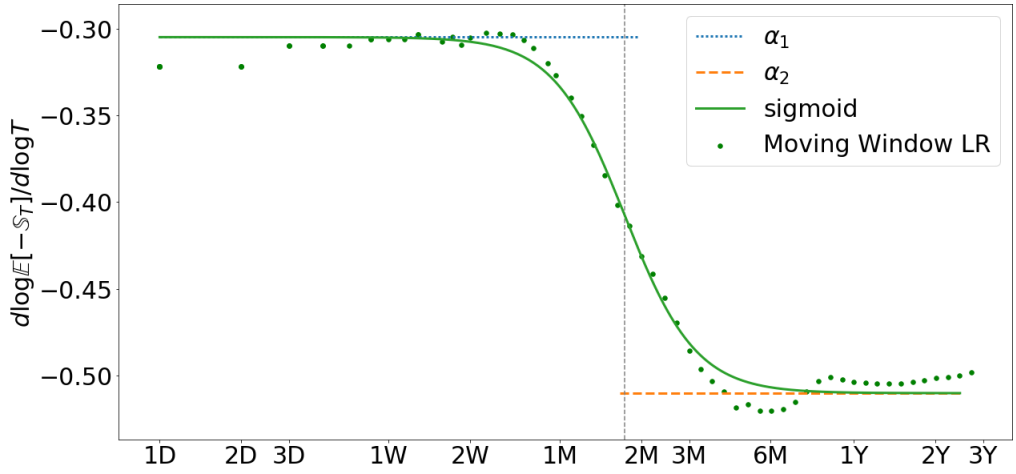

Figure 13: **An estimation of the log skew derivative** $\frac{d\log(-\mathbb{E}[\mathcal{S}_T])}{d\log(T)}$**.** We fit the data points for the average skew in the log-log plot in Figure 1 with linear regressions over a moving window. For each value of $\log(T)$ on the x-axis, the green dots represent the slopes of a linear regression over the moving window $[T_a, T_b]$ where $T_a = T \times e^{-1}$ and $T_b = \max(T \times e, 30 \text{ days})$, while the green curve is sigmoid fuction. The asymptotic values α_1 and α_2 are the slopes of the two linear regressions in Figure 2, performed separately for maturities below and above two months.

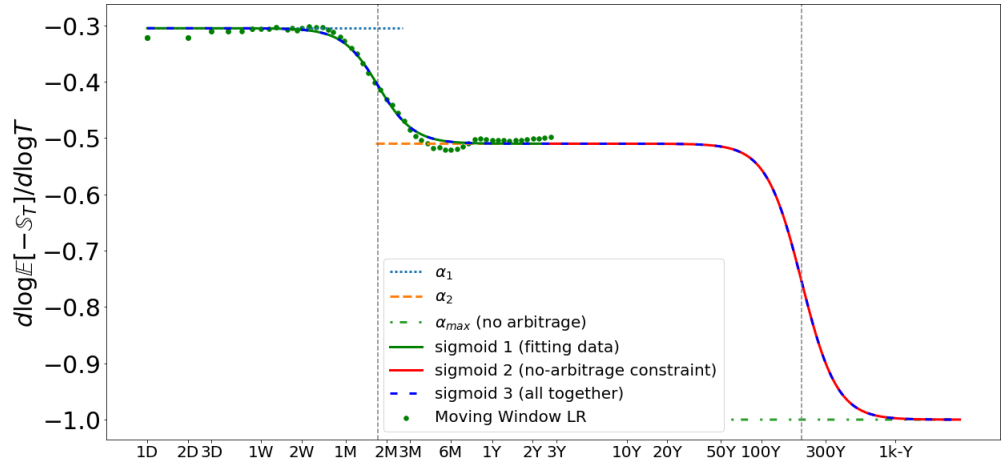

Figure 14: **Extrapolation of the log skew derivative consistent with no-arbitrage constraints.** We add a second sigmoid function of the form $(\alpha_{\max} - \alpha_2)\sigma\big(\kappa_2(\log T - \log T_2)$ to the model in Figure 13, with $\alpha_{\max} = -1$, $T_2 = 200$ years and $\kappa_2 = 3.3$. The resulting skew extrapolation, give in equation (14), is consistent with the no-arbitrage upper bound (10).

E Parameters time series

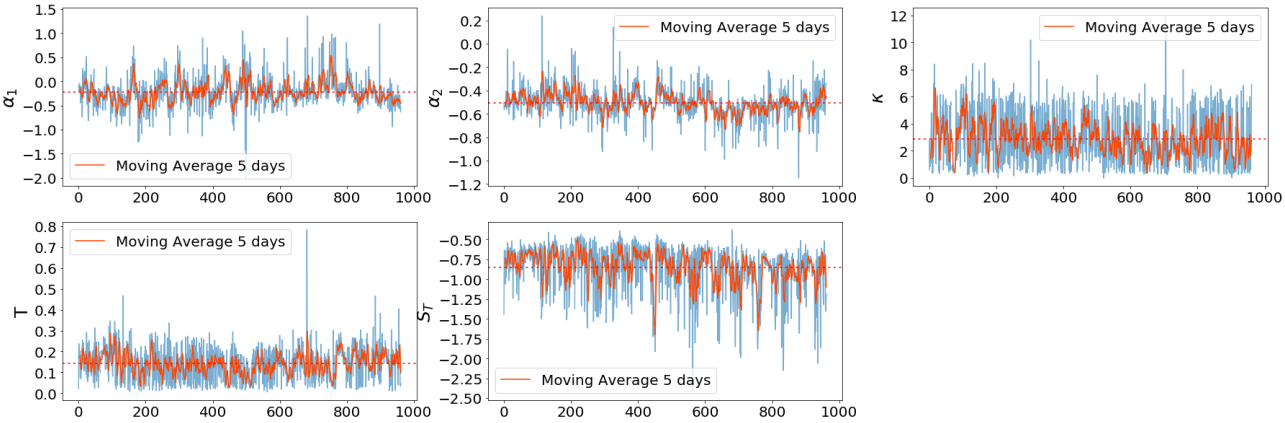

Figure 15: **Calibrated parameters for the 2PL5 model**. Daily time series of the parameters of the 2PL5 model fitted to the ATM skew of monthly and weekly options on every date in our data set between 2012 and 2015, using the regularized objective (9) with $\lambda = 10^{-3}$. The red dashed line shows the mean level of the whole series.

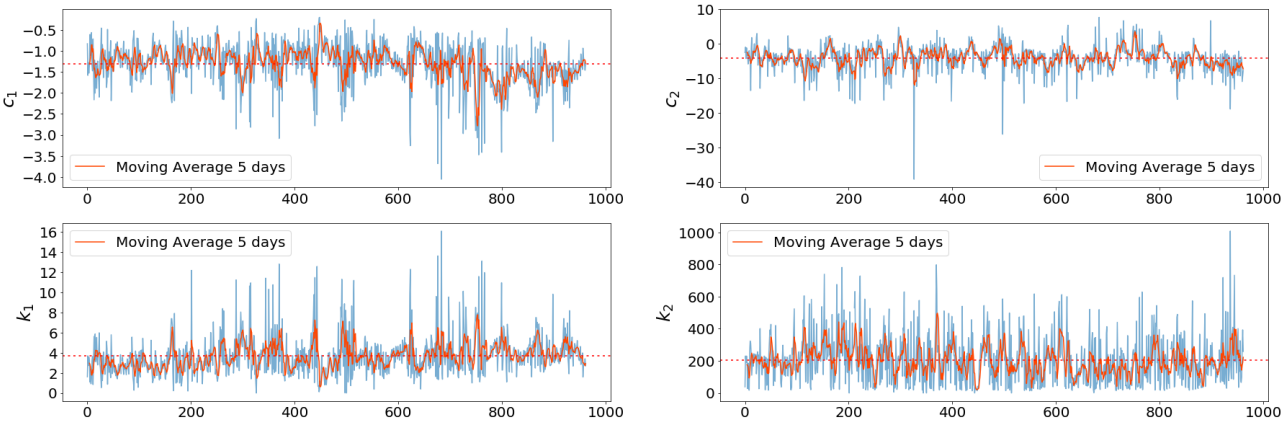

Figure 16: **Calibrated parameters for the 2fB model** . Daily time series of the parameters of the 2fB model fitted to the ATM skew of monthly and weekly options on every date in our data set between 2012 and 2015, using the regularized objective (9) with $\lambda = 10^{-3}$. The red dashed line shows the mean level of the whole series.