Path-Dependent Modeling and Forecasting of Implied Volatility Surfaces

Pablo Pastor, Guillem Ribas, Sergio Leal

April 10, 2025

Abstract

This report examines path-dependent modeling and forecasting techniques for implied volatility surfaces in financial markets. We investigate the impact of an underlying asset's historical values on at-the-money implied volatility by utilizing a model based on historical volatility and returns. To describe and calibrate a fit of the complete volatility surface, we consider two variations of the Stochastic Volatility Inspired (SVI) model, the Heston-like and Parsimonious SSVI. The Parsimonious SSVI model's enhanced fit to market data and decreased complexity make it especially successful. Through the combination of these frameworks, we are then able to support the claim that two parameters of the Parsimonious SSVI model are path-dependent, particularly when it comes to predicting the ATM region's development. Our results provide evidence for the existence of memory in implied volatility, providing a solid foundation for options market prediction modeling.

1 Introduction

Investors in the financial markets aim to make informed decisions to maximize their returns. In the realm of options trading, one of the most crucial metrics is the concept of implied volatility, allowing the measurement of the market's expectations of future price swings. Given its importance as a market metric, it is natural to consider a mathematical model to predict future implied volatilities for distinct maturities and option strike prices.

The purpose of [1], is to provide a model for the coherent forecasting of both implied volatility surfaces and the underlying asset returns in the market. To recall the definition of implied volatility, let $C(K,T)$ be the market price of a European call option with strike K and expiry T. Then, the implied volatility is defined as the unique parameter σ such that $C_{BS}(K,T,\sigma) = C(K,T)$, where C_{BS} denotes the Black-Scholes formula. We denote this implied volatility by $\sigma_{BS}(K,T)$. Usually, option prices are quoted in terms of their implied volatility since there is a one-to-one correspondence between the two, implying that volatility accounts for the same information as the price, and it allows to compare option prices on different underlying assets, given that it "scales" the price and eliminates the scale effect of the underlying's price. We then define the implied volatility surface (IVS) as the three-dimensional plot obtained by assigning $(K,T) \to \sigma_{BS}(K,T)$. If the Black-Scholes model exactly predicted market prices, one would expect to see a flat surface when plotting the volatility surface. However, it is well known that this is not the case, with volatility surface sheets exhibiting a convex structure. Furthermore, it is also observed from market data that the IVS is not fixed, and it evolves with time, inspiring the theory that past IVSs can be useful tools in forecasting future ones.

In Section 2, we begin by implementing the model introduced in [5], which attempts to estimate ATM implied volatilities using the path of past values of the underlying asset. To do so, we introduce the concept of at-the-money options. Having implemented this model, in Section 3 we then define a family of IVS parameterizations, attributed to [4]. Using these models, we then are able to calibrate their parameters for our entire dataset, defining a fitted IVS for every trading day in the data. Finally, we combine these models in 4, attempting to forecast the values of the model parameters using the path-dependent approach outlined in Section 2.

2 Path-dependent volatility

2.1 Guyon and Lekeufack's path-dependent volatility model

In this section, we study a model based on [5] to estimate at-the-money (ATM) implied volatility based on the past values of the underlying asset, hence showing that volatility is path-dependent. Let $(S_t)_{t\geq 0}$ be the price of an equity index and IV_t^{mkt} the ATM implied volatility, then the model writes

$$IV_t^{mkt} = \beta_0 + \beta_1 R_{1,t} + \beta_2 \Sigma_t. \tag{1}$$

The features $R_{1,t}$ and Σ_t are defined on a time grid $(t_i)_{i\in\mathbb{N}}$ as follows:

- $R_{1,t}$ is a trend feature given by:

$$R_{1,t} = \sum_{t_i \leq t} K_1(t - t_i) r_{t_i}, \tag{2}$$

  where $r_{t_i} = (S_{t_i} - S_{t_{i-1}})/S_{t_{i-1}}$ and $K_1 : \mathbb{R}_+ \to \mathbb{R}_+$ is a decreasing kernel weighting the past returns.

- Σ_t is an activity or volatility feature given by:

$$\Sigma_t = \sqrt{\sum_{t_i \leq t} K_2(t - t_i) r_{t_i}^2}, \tag{3}$$

  where K_2 is also a decreasing kernel.

The trend feature allows to capture the leverage effect, i.e the fact that volatility tends to rise when prices fall. This feature is a weighted average of past returns, giving more weight the returns from dates closer to t. This essentially captures, as its name indicates, the trend of the asset, in particular the direction and magnitude of the trend. Hence it allows to capture the leverage effect, noting that, typically, negative trends coincide with higher volatility periods. On the other hand, the volatility feature allows to capture volatility clustering, i.e. the fact that large changes in asset prices tend to be followed by large changes, and small changes tend to be followed by small changes. This is because the volatility feature tracks the evolution of past volatility, detecting high and low volatility periods, which allows the model to capture volatility clustering.

The Kernels used to assign weights to past data are assumed to follow a time-shifted power law (TSPL)

$$K_j(\tau) = \frac{Z_{\alpha_j,\delta_j}}{(\tau + \delta_j)^{\alpha_j}}, \quad j = 1, 2, \tag{4}$$

where Z_{α_j,δ_j} is a normalization constant such that $\sum_{t-C\leq t_i\leq t} K_j(t-t_i)\Delta = 1$, where $\Delta = 1/252$ is the business days frequency and C is a hyperparameter (cut-off lag) controlling at which point the sums in R_1 and Σ are truncated, thus, choosing the amount of previous days used to predict the next implied volatility value.

In order to measure the effectiveness of the path-dependent volatility (PDV) model to explain the evolution of implied volatility, the R^2 score is used as metric, which has the following expression

$$R^2(y,\hat{y}) = 1 - \frac{\sum_{i=1}^{n}(y_i-\hat{y}_i)^2}{\sum_{i=1}^{n}(y_i-\bar{y}_i)^2}, \tag{5}$$

where $y = (y_i)_{1\leq i\leq n}$ are the observed data points, $\hat{y} = (\hat{y}_i)_{1\leq i\leq n}$ are the predicted data points and we define $\bar{y}_n = \frac{1}{n}\sum_{i=1}^{n} y_i$.

2.2 Data treatment

We consider data from call and put options on the S&P 500 for all business days between 2003 and 2023. The data includes implied volatilities and bid and ask prices for several maturities and strikes. The mid price is then obtained as an average between bid and ask prices, and we import the underlying price for each date. We find that there is way more data for recent years and, for each year, the majority of the data corresponds to short maturities.

The main issue we face with the data is finding the ATM volatility. In order to do that, we make use of the put-call parity, which we recall here

$$C(K,T) - P(K,T) = S - DK,$$

where D is the discount factor. This equality holds because a portfolio made of one long call option and one short put option with the same strike and expiry will always be worth $S-K$ at expiry, regardless of whether the underlying asset ends up being above or below the strike. Hence, by no arbitrage, the same equality must hold at all previous times. If an option is ATM, then by definition it must hold that $S - DK = 0$, so the strike is equal to the forward price of the asset. Therefore, our strategy will be to try to find, for each T, the value of K for which $C(K,T) = P(K,T)$. However, usually we won't be able to find such values in our data. The approach will then be to run a linear regression, for each T, of $C(K,T) - P(K,T)$ against K

$$C(K,T) - P(K,T) = \alpha K + \beta,$$

the value of K representing the ATM level will be the one for which

$$\alpha K_{ATM} + \beta = 0 \Rightarrow K_{ATM} = -\frac{\beta}{\alpha}.$$

The main issue we found when trying to implement this procedure is not having enough data for all maturities. If we want the linear regression approach to work, we need to consider just strikes that are close to the ATM strike. As we don't know the ATM strike yet, we use the spot as benchmark to decide whether a strike is close to the ATM strike and hence use it in the linear regression. However, when applying this filter, we notice some maturities with no strikes close enough to the spot, and hence we are not able to run the linear regression. Even for the cases with just one data point, if we apply the linear regression, the result will be a flat line with slope equal to 0, and we will not obtain a reasonable forward price. Therefore, we run the

regression for maturities with at least two strikes that are close enough to the spot.

After applying the treatment of the data, we have reduced the size of our dataset, and, considering we already had few data points for longer maturities, our forecast of implied volatility for longer maturities is less accurate than for shorter maturities, where our results show the model is able to reproduce the behaviour of implied volatility.

Given that we want to measure the out-of-sample performance of the model, we start by splitting our data into two sets, a train set and a test set. The train set ranges from March 8, 2012 to December 31, 2020 and the test set goes from January 1, 2021 to December 30, 2022.

2.3 Calibration and results

The PDV model with a TSPL kernel, depends on seven key parameters: $(\alpha_1, \delta_1, \alpha_2, \delta_2)$, corresponding to the two TSPL kernels K_1 and K_2, and $(\beta_0, \beta_1, \beta_2)$, which represent the intercept, the impact of the trend signal, and the influence of volatility, respectively. These parameters are tuned individually for each maturity through a two-stage calibration. Initially, a set of EWMAs of the underlying returns and squared returns (spanning 10 to 250 days) are used in ridge regressions to estimate preliminary values for (α_1, δ_1) and (α_2, δ_2) via the TSPL kernels K_1 and K_2, respectively. These estimates are then held fixed while a linear regression on the ATM implied volatility, using the features R_1 and Σ, provides initial values for $(\beta_0, \beta_1, \beta_2)$. Finally, starting from these initial values, the seven parameters are jointly optimized by minimizing the objective below using *scipy*'s *least_squares* function [3] with the trust-region reflective method, defined as

$$\begin{aligned} &\min_{(\alpha_1,\delta_1,\alpha_2,\delta_2,\beta_0,\beta_1,\beta_2)\in\mathbb{R}^7} \sum_{t\in\mathcal{T}_{\text{train}}} \left(IV_t^{\text{mkt}} - \beta_0 - \beta_1 R_{1,t} - \beta_2 \Sigma_t\right)^2 \\ &\text{s.t.} \quad \alpha_j, \delta_j \geq 0 \quad \text{for } j \in \{1,2\}, \\ &\qquad R_{1,t} = \sum_{t-C\leq t_i\leq t} \frac{Z_{\alpha_1,\delta_1}}{(t-t_i+\delta_1)^{\alpha_1}} r_{t_i}, \\ &\qquad \Sigma_t = \sqrt{\sum_{t-C\leq t_i\leq t} \frac{Z_{\alpha_2,\delta_2}}{(t-t_i+\delta_2)^{\alpha_2}} r_{t_i}^2}, \end{aligned} \tag{6}$$

where $\mathcal{T}_{\text{train}}$ is the set of dates in the train set, IV_t^{mkt} is the market ATM implied volatility observed at time t for some fixed maturity and C is a cut-off lag.

We use the algorithm described in [7] as a baseline and we adapt it for our particular problem. In order to truncate the sums in the computations of R_1 and Σ, we used a lag of 800 for maturities shorter than 6 months and 1,000 for longer maturities, after a hyperparameter tuning on a validation set using a grid search in {200, 400, 600, 700, 800, 900, 950, 1000, 1050, 1100, 1200, 1500}. As we can see on Figure 1, the R^2 scores on the test set drop after $T = 4$, and this is due to the lack of data after those maturities, where for $T = 7$, the algorithm does not have enough data for the estimations. This is shown in Figure 2, where it can be seen that the first 4 maturities (1-4 months) accumulate more than 75% of the data, leading to poor results for longer maturities.

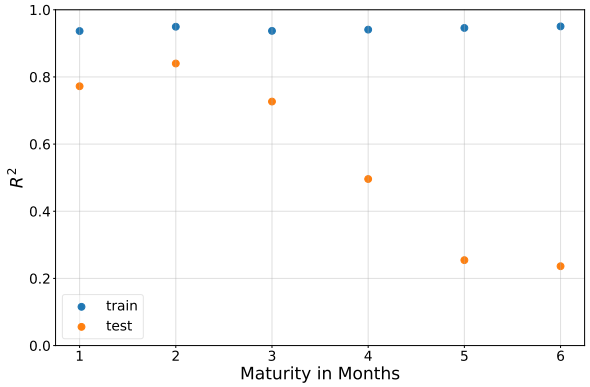

Figure 1: R^2 scores on the train and the test sets as a function of the ATM implied volatility maturity for $C = 800$ when $T < 6$ months and $C = 1000$ otherwise.

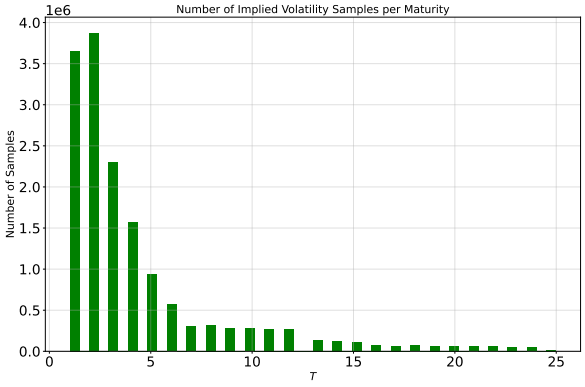

Figure 2: Number of implied volatility samples per maturity in the data.

The reality is that even though the number of years used for training and testing is 80% and 20% respectively, as highlighted in [1], this is not reflected in the quantity of data available in the dataset, where for these dates, the test set includes 40% of the data, something not recommendable for most models. To address this issue, we adjust the training and testing periods to achieve a true 80-20 chronological split. The new test set spans from March 23, 2022, to August 31, 2023. Figure 3 displays the 1-month ATM implied volatility alongside the underlying asset price over both the training and test sets. The hyperparameters are calibrated using data from September 23, 2021, to March 23, 2022. The latest R^2 scores obtained from the model are shown in Figure 4.

With this adjustment, we now have sufficient data to train and evaluate models for maturities up to 12 months. The performance on the test set is notably more consistent attributing the improvement to the inclusion of data from 2021 in the training period, which shares similar implied volatility levels and data density with 2022 and 2023.

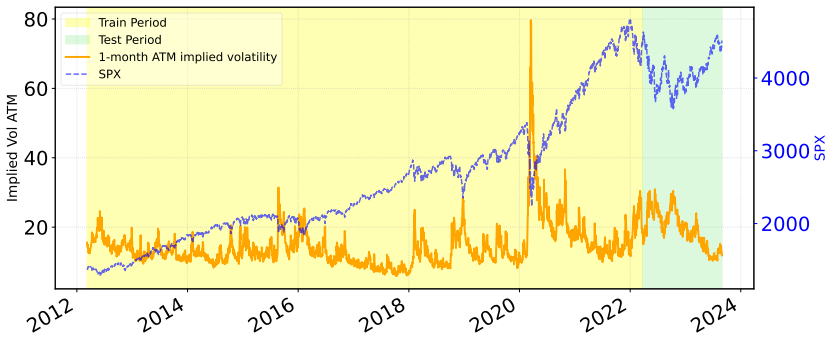

Figure 3: Joint evolution of the 1-month ATM implied volatility and its underlying index from March 8, 2012 to August 31, 2023. The split between the train and the test sets is represented through the use of different background colors.

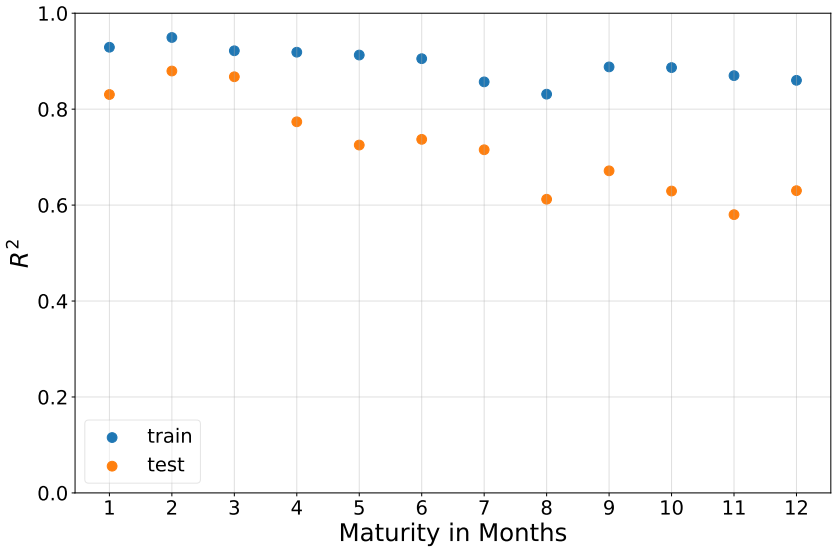

Figure 4: R^2 scores on the train and the test sets as a function of the ATM implied volatility maturity for $C = 800$ when $T < 6$ months and $C = 1000$ otherwise.

Considering the absence of results for maturities longer than 12 months due to lack of sufficient data, the results are very similar to those presented in [1], although we observe a slightly smaller deterioration. This is mainly because, in our case, the parameters are calibrated using a validation set that more closely resembles the test set than the initial part of the training set, resulting in more accurate estimates. The results for the remaining maturities are also quite similar, except for $T = 8$ and $T = 11$, where the test scores drop slightly below 60%. Figure 5 shows the predicted implied volatilities compared to the observed ones for $T = 1$ in a more detailed manner, including the underlying price.

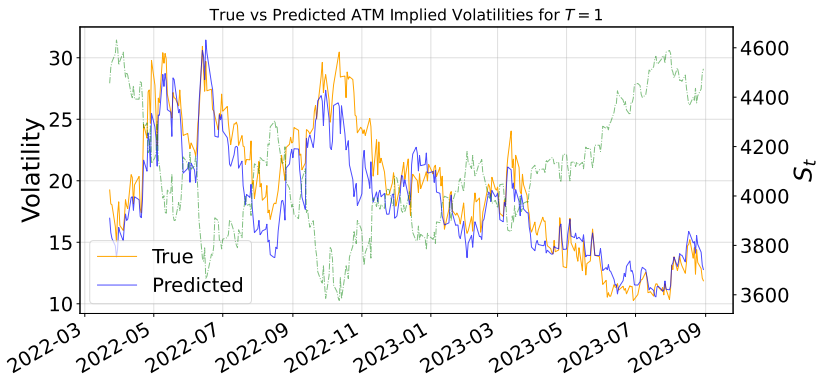

Figure 5: Comparison of predicted and true ATM implied volatilities for $T = 1$ compared to the index price.

Another fact to comment on is the distance between the train set R^2 values and the test ones, a symptom of overfitting of the model. As commented in [1], this is not due to the complexity of the model, but rather, due to the fact that the implied volatilities are much more different during the test set than during the train set. Events such as war, inflation, the rise of interest rates, etc. have affected the extent to which the volatility reacts to the underlying index

movements. This is depicted in Figure 3, where we can see that during the train set, except for a crash in March 2020, the S&P 500 has experienced a constant increase with very little variations on the train set while the test set is characterized by a market with high volatility.

In order to avoid this overfitting, we include a regularization parameter λ to penalize large values of the kernel parameters $\alpha_1, \alpha_2, \delta_1, \delta_2$ during the calibration. This regularization term discourages the model from assigning excessively low weights to distant past returns and/or squared returns. This ensures that the memory of the model does not decay too rapidly, preserving the influence of older observations. This is done by introducing an L^2 penalization term in the objective function so that Equation 6 becomes:

$$\min_{(\alpha_1,\delta_1,\alpha_2,\delta_2,\beta_0,\beta_1,\beta_2)\in\mathbb{R}^7} \sum_{t\in\mathcal{T}_{\text{train}}} \left(IV_t^{mkt} - \beta_0 - \beta_1 R_{1,t} - \beta_2 \Sigma_t\right)^2 + \lambda \left(\sum_{j=1}^{2} \alpha_j^2 + \sum_{j=1}^{2} \delta_j^2\right),$$
$$\text{s.t.} \quad \alpha_j, \delta_j \geq 0 \quad \text{for } j \in \{1, 2\}.$$

We evaluate the R^2 scores for $\lambda \in \{10^{-2}, 10^{-3}, \ldots, 10^{-7}\}$. Figure 6 shows the R^2 scores on the test set for the different values of λ. We observe that for higher values of λ, there are noticeable drops in performance. The optimal value, considering the average performance on both the train and test sets, is $\lambda = 10^{-4}$, although the scores are quite similar for the lowest values. The results are also very similar to those obtained with $\lambda = 0$, and therefore we conclude that regularization does not have a significant enough effect to mitigate the overfitting problem.

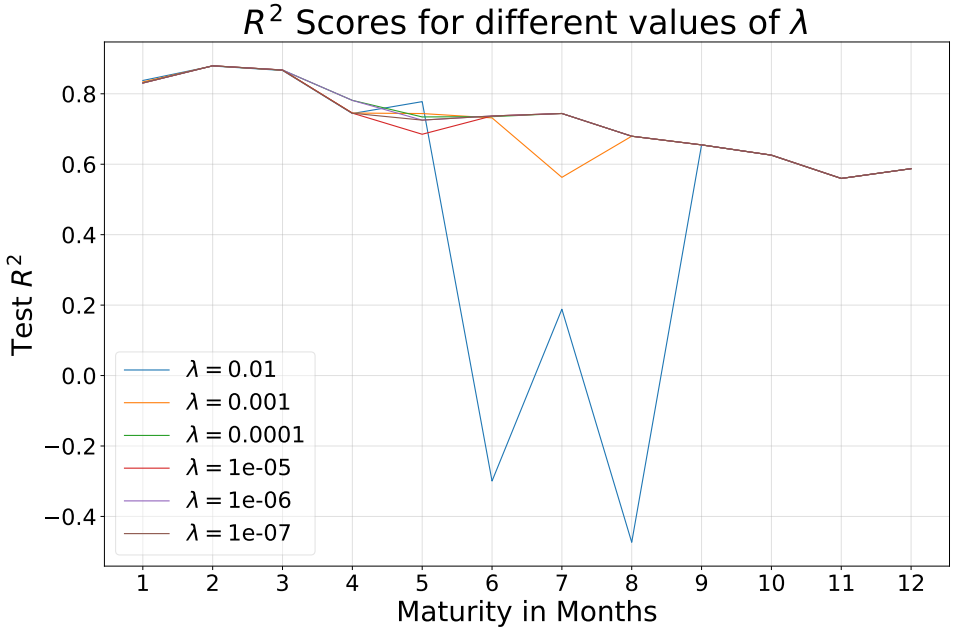

Figure 6: R^2 scores in the test set for different values of the regularization parameter λ.

Beyond this quantitative comparison, the empirical analysis in this section shows that the historical path of the underlying asset continues to influence the ATM implied volatility. Although this effect becomes weaker as maturity increases, it still plays an important role even for the longest maturities we consider. Also, the impact of the underlying asset on implied volatility seems to have a long memory - to predict the ATM implied volatility for longer maturities we use

almost 4 years of past data. This result highlights the importance of taking long-term patterns into account when modeling the term structure of implied volatility, especially for longer-dated options.

3 SSVI and calibration

In this section, we will introduce the Stochastic Volatility Inspired (SVI) and Surface SVI (SSVI) models for the implied volatility surface of options in the market. Under these frameworks, we will then be able to calibrate the parameters in these models to fit the market implied volatilities, which will in turn pave the way to explore their path-dependency in Section 4.

3.1 Arbitrage and the SSVI model

When modeling the implied volatility surfaces, we want to ensure that they remain free from static arbitrage, that is, there should be no arbitrage opportunity by simply trading options with the prices determined by the Black-Scholes formula under their implied volatilities.

In [1], we are given a set of conditions that prove sufficient for the implied volatility surface to be free of static arbitrage. We state them here without proof to illustrate the concept of calendar arbitrage, needed for later.

Theorem 1 (Non-arbitrage conditions)**.** *Define the total implied variance as* $\omega(k,T) = T\sigma_{BS}^2(k,T)$*, where* $\sigma_{BS}(k,T)$ *is the Black-Scholes implied volatility associated with the log-strike* k *and maturity* T*. Note that we define* $k = \log\left(\frac{K}{S_0 e^{rT}}\right)$*, where* K *is the strike of the option,* S_0 *is the spot price of the underlying asset and* r *is the risk-free interest rate. Then, if* $\omega : \mathbb{R} \times \mathbb{R}_+ \to \mathbb{R}_+$ *satisfies*

(i) $\omega(\cdot, T)$ *is of class* C^2 *for all* $T \geq 0$*,*

(ii) $\omega(k,T) > 0, \quad \forall (k,T) \in \mathbb{R} \times \mathbb{R}_+^*$*,*

(iii) for all $(k,T) \in \mathbb{R} \times \mathbb{R}_+^*$ *we have*

$$\left(1 - \frac{k\partial_k \omega(k,T)}{2\omega(k,T)}\right)^2 - \frac{\partial_k \omega(k,T)^2}{4}\left(\frac{1}{\omega(k,T)} + \frac{1}{4}\right) + \frac{\partial_{kk}^2 \omega(k,T)}{2} \geq 0,$$

(iv) $\omega(k,\cdot)$ *is non-decreasing* $\forall k \in \mathbb{R}$*,*

(v) for all $T > 0$ *we have that*

$$\lim_{k \to +\infty} \frac{-k}{\sqrt{\omega(k,T)}} + \frac{\sqrt{\omega(k,T)}}{2} = -\infty, \quad \textit{and}$$

(vi) $\omega(k,0) = 0 \quad \forall k \in \mathbb{R}$*,*

then the implied volatility surface described by ω *is free of static arbitrage.*

With these conditions in mind, due to [4] it is said that an implied volatility surface is free of calendar arbitrage if condition *(iv)* is satisfied. Additionally, an implied volatility surface is also thought of being free from butterfly arbitrage if conditions *(iii)* and *(v)* hold.

Originally introduced in 1999 at Merill Lynch, the SVI model is a parameterization of a slice for a given maturity of the implied volatility surface.

Definition 2 (Stochastic Volatility Inspired)**.** For a given maturity $T > 0$, the SVI parameterization is given by

$$\omega(k,T) = a_T + b_T\left(\rho_T(k-m_T) + \sqrt{(k-m_T)^2+\sigma_T^2}\right)$$

with $a_T \in \mathbb{R}, b_T \geq 0, |\rho_T| < 1, m_T \in \mathbb{R}$ and $\sigma_T > 0$. Additionally, the parameters must satisfy $a_T + b_T\sigma_T\sqrt{1-\rho_T^2} \geq 0$ in order for ω to remain positive $\forall k \in \mathbb{R}$.

Though this model is highly tractable and can fit the market volatilities well, two main drawbacks exist. Firstly, as previously mentioned, the model is a parameterization of a single slice of the implied volatility surface, with the map $k \mapsto \omega(k,T)$ for a given maturity T, due to all of the parameters being dependent on the value of T. This makes the fitting of an entire surface a lot more tedious as distinct maturities have to be individually fitted, making it unable to generate surface sheets for intermediate values of T that aren't present in the calibration data. Moreover, conditions on the parameters to ensure no butterfly arbitrage are very convoluted. For these reasons, [4] introduces the SSVI model as an extension to the original SVI parameterization to solve these problems.

Definition 3 (Surface SVI)**.** Let $\varphi : \mathbb{R}_+^* \mapsto \mathbb{R}_+^*$ be a smooth function such that $\lim_{T\to 0}\theta_T\varphi(\theta_T)$ exists in $\mathbb{R}$, where we have that $\theta_T := \sigma_{BS}^2(0,T)T$ is the ATM total implied variance. The SSVI implied volatility surface is defined by

$$\omega(k,T) = \frac{\theta_T}{2}\left(1 + \rho\varphi(\theta_T)k + \sqrt{(\varphi(\theta_T)k+\rho)^2 + (1-\rho^2)}\right).$$

The power behind this model lies in the fact that, while the SVI model uses 5 parameters to characterize a single slice of the implied volatility surface, the SSVI parameterization only requires the choice for the function φ and $\rho \in (-1,1)$, with a set of values θ_T, for which we can use the ATM market implied volatilities as an initial guess for calibration, to model the entire surface. This is because, while it seems that we encounter the same problem as with the SVI parameterization, with only single sheets being fitted for each maturity, the calibrated θ_T parameters can be interpolated to generate sheets for intermediate maturities.

The choice of φ is crucial to ensure that the SSVI fits the market data well whilst still being free from static arbitrage. In [1], three distinct choices for φ are presented, however, in this report, we will only focus on one, namely

$$\varphi(\theta) := \frac{1}{\lambda\theta}\left(1 - \frac{1-e^{-\lambda\theta}}{\lambda\theta}\right), \quad \lambda > 0,$$

known as the Heston-like parameterization.

Remark 4 (Heston-like SSVI arbitrage conditions)**.** *The Heston-like SSVI surface is free of calendar arbitrage if the following are satisfied*

(i) $\partial_T\theta_T \geq 0, \quad \forall T > 0,$

(ii) $\lambda \geq \frac{(1+|\rho|)}{4}$.

3.2 Calibrating the Heston-like SSVI model

With the tools introduced in Section 3.1, we can now use real market data to replicate historical implied volatility surfaces. It is well-known that the daily implied volatility surfaces recovered from data providers may contain arbitrage due to interpolations of market quotes, hence, we first proceed to check for arbitrage in our dataset. Recalling condition *(iv)* in Theorem 1, we want our market total variance surface sheets, $\omega(k, T_i)$ for $T_i \in \mathcal{T}$ where $\mathcal{T}$ is the set of option maturities in the dataset and $i = 1, \ldots, |\mathcal{T}|$, to satisfy $\omega(k, T_1) \leq \omega(k, T_2)$ for $T_1 < T_2$, in order to ensure the absence of calendar arbitrage.

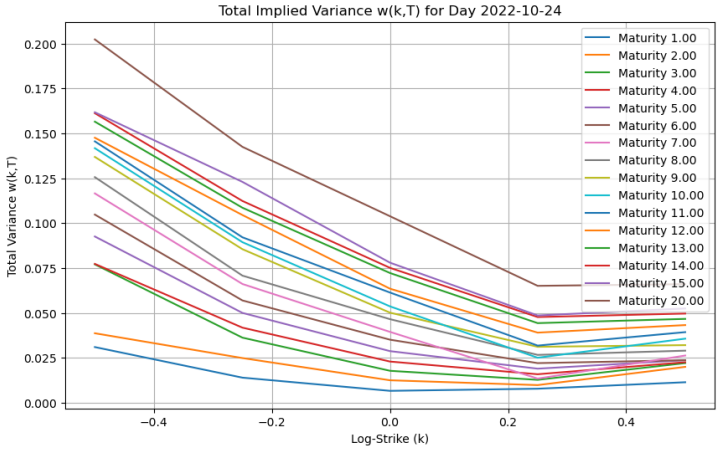

Figure 7: Market total variance surface sheets for 2022-10-24 for each available maturity, T, in months.

Figure 7 is an example of a day in the dataset containing calendar arbitrage. This can be seen clearly by noting that, if condition *(iv)* in Theorem 1 were to be satisfied, no crossings would occur between the line representing each sheet. In treating the data, [1] removes any day in the dataset that has any occurrence of these total variance sheet crossings in order to calibrate the models with arbitrage-free data. In the case of our dataset, we noted that more than 75% of the days contain at least one of these crossings, indicating the presence of calendar arbitrage. Clearly, this proves to be very limiting in calibrating the Heston-like SSVI parameterization, as we require arbitrage-free data for the calibration. Nevertheless, we noted that on more than 95% of the data points, the sheets for $k \in (-0.1, 0.1)$ were free from calendar arbitrage. This led us to remove only those days where arbitrage was present in this subset of the surface, ensuring a good fit for ATM implied volatilities.

To calibrate the parameterization, we use the data for each day in the dataset, noting that each day gives rise to an implied volatility surface. To do so, we follow the same approach as in [1], by utilizing the `minimize` function using the Sequential Least Squares Programming (SLSQP) [3] algorithm from the SCIPY package in Python [2]. The objective function to minimize is

defined as

$$\min_{(\theta_{T_i})_{i=1,\dots,|\mathcal{T}|},\rho,\lambda} \sum_{i=1}^{|\mathcal{T}|} \sum_{k\in\mathcal{K}_{T_i}} \phi(k)(\sigma_{\text{Market}}(k,T_i) - \sigma_{\text{SSVI}}(k,T_i,\theta_{T_i},\lambda,\rho))^2$$

$$\begin{aligned} \text{s.t.} \quad & \theta_{T_1} \geq 0, \\ & \theta_{T_{i+1}} \geq \theta_{T_i} \quad \text{for } 1 \leq i \leq |\mathcal{T}| - 1, \\ & (\rho,\lambda) \in (-1,1) \times \mathbb{R}_+^*, \\ & \lambda \geq (1+|\rho|)/4, \end{aligned}$$

where $\mathcal{K}_T$ is the set of log-strikes for a given maturity T, $\phi(k)$ is the standard normal density function calculated at k, σ_{Market} is the market implied volatility and σ_{SSVI} is the implied volatility yielded by the Heston-like SSVI parameterization under the given parameters.

$(\theta_{T_i})_{i=1,\dots,\lvert\mathcal{T}\rvert}$	ρ	λ
$(\sigma_{\text{Market}}(0,T_i)^2 T_i)_{i=1,\dots,\lvert\mathcal{T}\rvert}$	-0.5	1

Table 1: Initial values for Heston-like SSVI calibration.

In following the initial guesses used in [1], summarized in Table 1, we encounter one of the major difficulties in calibrating the model using our data. This is because, firstly, $\sigma_{\text{Market}}(0,T)$ may not exist in the dataset, forcing us to estimate it through linear interpolation of existing market volatilities at the given maturity T. Secondly, the initial guesses θ_{T_i} for $i = 1,\dots,|\mathcal{T}|$, should already satisfy the constraints set in the optimization problem in order to be in the feasible region of the optimization problem. However, since the dataset contains mostly days where calendar arbitrage is present for $|k| > 0.1$, as exemplified in Figure 7, the second constraint is not satisfied for most days, namely, there exist some maturities T_i and T_j such that $T_i < T_j$ and $\theta_{T_i} > \theta_{T_j}$ for $i,j = 1,\dots,|\mathcal{T}|$. We recognize that there exist numerous techniques to remove arbitrage from these datasets, as outlined in [1], however, this falls outside of the scope of the report and is left for further work. As a result, the presence of arbitrage in the initial guess poses a lot of difficulties for the SLSQP algorithm, as convergence to a solution from an infeasible initial point is not only much slower but also not guaranteed [6].
In Figure 8, we can see how the Heston-like SSVI parameterization achieves a vaguely similar implied total variance surface to the market on 2022-10-24, mimicking the overall downward trend as k increases and achieving a 17.28% mean percentage error. We can partly explain the error by recalling the existence of calendar arbitrage for values of k where $|k| > 0.1$, as the Heston-like SSVI model calibration may have begun with an initial guess outside of the feasible region. Additionally, the error may also be explained by the fact that the function φ is bounded above by 0.5, restricting the shape that the surface can take.

The large number of parameters of this model, which noticeably slows down the optimization problem, as well as the inability to achieve a satisfactory calibration using a dataset where arbitrage is present, motivate the introduction of a more parsimonious model requiring fewer parameters to be calibrated.

3.3 The Parsimonious SSVI model

Stemming from the need of an SSVI parameterization that requires less parameters to be calibrated, the Parsimonious SSVI model is introduced in [1].

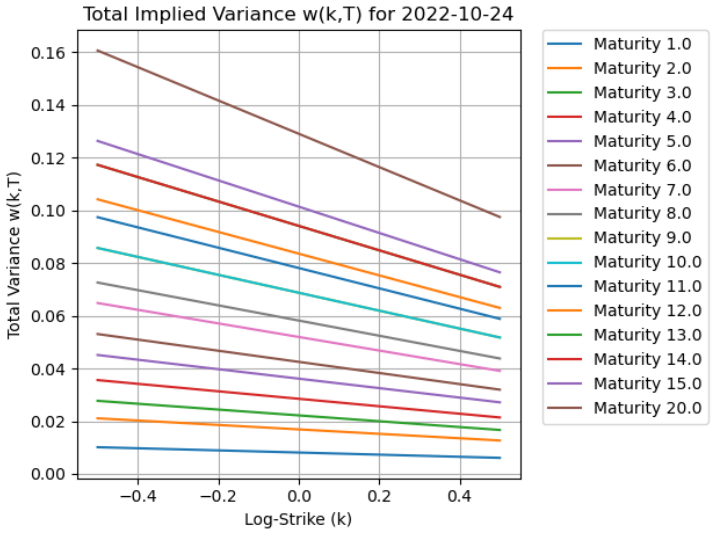

Figure 8: Calibrated Heston-like SSVI total variance surface sheets for 2022-10-24 for each maturity present in the dataset, T, in months.

Definition 5 (Parsimonious SSVI)**.** For a given maturity T and log-strike k, the Parsimonious SSVI implied volatility surface is parameterized as

$$\omega(k,T) = \frac{\theta_T}{2}\left(1 + \rho\varphi(\theta_T)k + \sqrt{(\varphi(\theta_T)k+\rho)^2 + (1-\rho^2)}\right),$$

where $\varphi(\theta) = \frac{\eta}{\sqrt{\theta(1+\theta)}}$ and $\theta_T = aT^p$ such that $\rho \in (-1,1)$, $a, p \geq 0$ and $\eta > 0$.

Clearly, the number of parameters in this model is more convenient for the calibration in larger datasets, remaining fixed at 4 for any given number of maturities. In addition, however, the parameterization of the values θ_T, is much more tractable in our dataset, as the initial guess of the minimization problem no longer imposes constraints on θ_{T_i} for $i = 1, \ldots, |\mathcal{T}|$, allowing for the successful calibration of the parameters using the same methodology. In this case, the objective function for the minimization problem is the same as the one used in Section 3.2, in the parameter space spanned by (a, p, ρ, η) where the only constraint is given by

$$(a, p, \rho, \eta) \in \mathbb{R}_+ \times \mathbb{R}_+ \times (-1,1) \times \mathbb{R}_+^*.$$

Furthermore, given that our dataset is missing information for many maturities, this parameterization has the added benefit that implied total variance sheets for missing maturities can be simulated without interpolating, as the parameters do not depend on T.

a	p	ρ	η
0.5	1	-0.5	1

Table 2: Initial values for Parsimonious SSVI calibration.

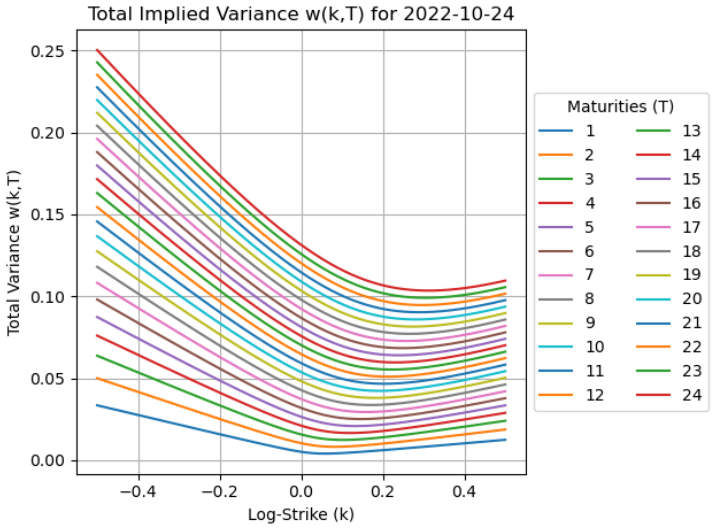

Figure 9: Market total variance surface sheets for 2022-10-24 for all maturities $T = 1, \ldots, 24$ in months.

Resulting from the calibration of the Parsimonious SSVI model, the implied total variance surface shown in Figure 9 is a considerable improvement from the Heston-like parameterization, more closely resembling the market surface with a 6.51% mean percentage error. We attribute this improvement to two crucial differences. First, the parameterization of θ_T ensures that the initial guess is always within the feasible region of the problem. Secondly, the function $\varphi(\theta)$ for the Heston-like parameterization is bounded above by 0.5, whilst there is no such restriction in the case of the Parsimonious SSVI, which imposes fewer constraints on the shape of the surface of the latter, achieving a better fit [1].

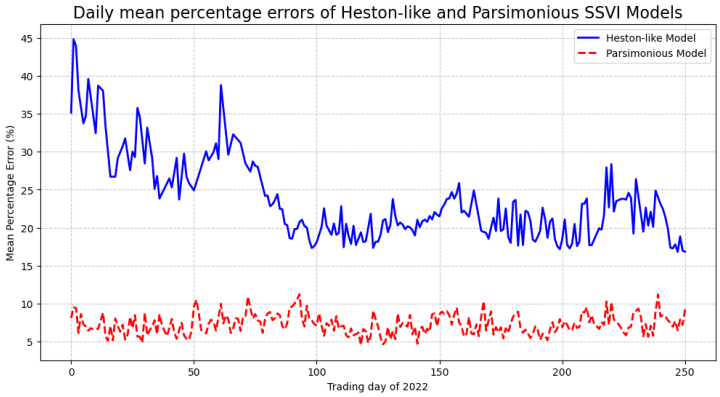

Figure 10: Daily mean percentage errors of Heston-like and Parsimonious SSVI model throughout the trading days of 2022.

Figure 10 illustrates the ability of the Parsimonious SSVI to consistently provide a good fit to the implied volatility surfaces in the market. The small fixed number of parameters and its consistency in accuracy then justify the choice of using this model for the path-dependent volatility surface calibration in Section 4.

4 Path-dependent volatility surface

The idea of this final section is to combine the results obtained in Sections 2 and 3 to forecast the evolution of volatility surfaces. In Section 2, we observed that the evolution of the ATM implied volatility could be explained using the past evolution of the underlying index. Furthermore, in Section 3, we used a parsimonious version of the SSVI parametrization to fit volatility surfaces. We would like to combine both approaches and be able to forecast volatility surfaces based on the past values of the underlying index. In order to achieve this, the idea is to model each of the 4 parameters of the parsimonious SSVI model as a function of the underlying index's path. Hence, we try to fit the path-dependent volatility model presented in Section 2 to the evolution of each of the 4 parameters of the parsimonious SSVI obtained from the calibration in Section 3.

4.1 No ATM filtering

In [1], the authors mention they are able to obtain good results just for parameters a and p, while for ρ and η the R^2 scores are not good enough. This is to be expected, as in Section 2, we were able to explain the ATM variance using the past values of the underlying index, and the two parameters of the Parsimonious SSVI that try to reproduce the ATM total variance are a and p.

The results obtained for both the training and test sets, along with the hyperparameters used, are presented in Table 3. These results are good for a, which indicates that we are able to explain a significant part of the volatility surface evolution using past values of the underlying index. In contrast, the results for p are noticeably poorer, which was somewhat expected given the nature of this parameter.
Regarding p, [1] focuses on fitting $\log(p)$ instead of p, given that the calibration results are more satisfactory. This is likely because the part of the Parsimonious SSVI parametrization representing the ATM total variance is $\theta_T = aT^p$, hence the dependence on p is exponential. However, using $\log(p)$ allows for negative values, whereas our original model from Section 2 was designed for implied volatilities, which must remain positive. To address this, we decided to shift $\log(p)$ by 0.2 (constant large enough to avoid negative values), and then undo the transformation after calibration.

For the calibration, we split the data into training and test sets, which span from 08/03/2012 to 28/02/2022 and from 01/03/2022 to 01/01/2023, respectively. For the cut-off lag C, we used the same value for both the trend and volatility features.

Parameter	**(C, λ)**	**Train**	**Test**
a	$(700, 10^{-4})$	66.7%	55.2%
p	$(2000, 10^{-2})$	27.4%	30.3%

Table 3: R^2 scores obtained for the daily evolution of a and p.

The authors mention that obtaining good results for p was an unexpected finding since, even though p contributes to explain the ATM total variance, it mostly explains how ATM variance depends on expiry, and it is less clear how past values of the underlying index can affect this dependence, since it is not homogeneous over all maturities.

We represent the test-set true and predicted values for p along with the stock path in Figure 12 to show how our model is able to forecast the evolution of p and how this evolution is correlated with the path of the underlying index.

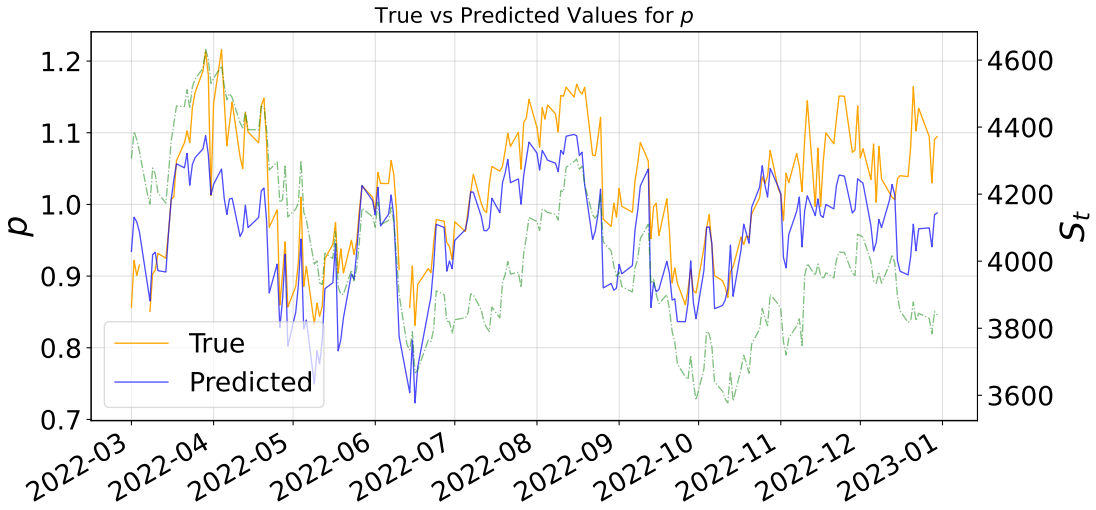

Figure 11: Stock price, predicted and true p values.

4.2 ATM Filtering

In an attempt to improve the results, we apply the procedure described in Section 3.2, filtering the data to retain only ATM observations and repeating the experiments, that is, we just consider options whose log-strike absolute value is less than 0.1 (i.e $|k| < 0.1$) for the daily calibration of the Parsimonious SSVI, ensuring a and p are properly calibrated. This is a natural choice, as the parameters a and p parametrize the ATM total variance. Table 4 presents the results, where an improvement can be observed, particularly in the estimation of p, as both the training and test scores are higher. This is very surprising given the nature of the parameter and the comment made by the authors mentioned before. The values for a remain very similar to those obtained previously — slightly lower in the training set and unchanged in the test set.

Parameter	**(C, λ)**	**Train**	**Test**
a	$(800, 10^{-4})$	60.9%	54.9%
p	$(2000, 10^{-2})$	68.8%	61.5%

Table 4: R^2 scores obtained for the daily evolution of a and p computed with ATM data.

Finally, Figure 4 shows the accuracy of the prediction of p on the full dataset. As can be seen, the predictions are highly accurate, especially after 2018, even capturing the drop in March 2020.

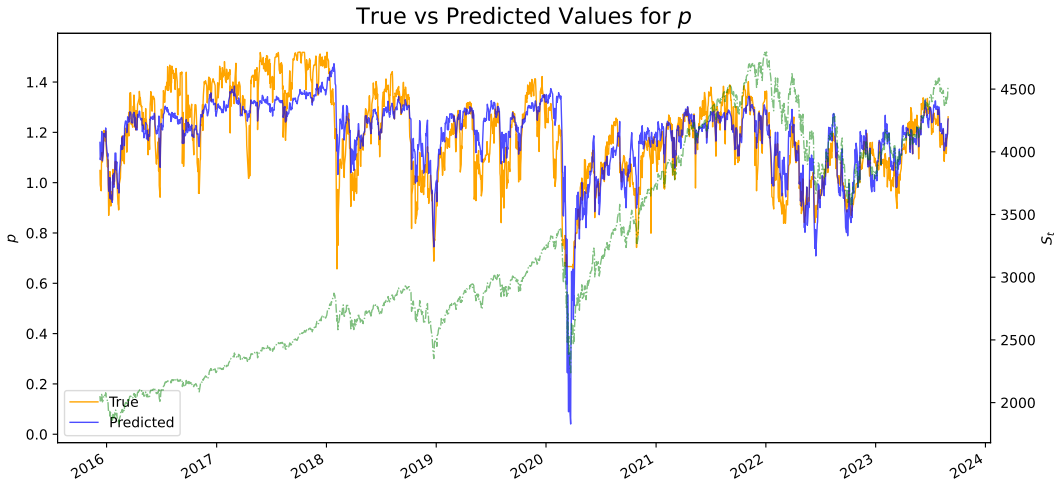

Figure 12: Stock price, predicted and true p values computed with ATM data. The green dashed line represents the path of the SP500 index.

5 Conclusions

In this project, we explored how the past behavior of an underlying index can help explain the evolution of the implied volatility surface, focusing on at-the-money options. Building on existing path-dependent models, we confirmed that two key features, the weighted averages of past returns and past squared returns, contain useful information for predicting implied volatility.

In exploring parameterizations of the implied volatility surface, we introduced two variants of the SSVI model, most notably, the Parsimonious SSVI parameterization, which describes the implied volatility surface using only four parameters. This version is particularly effective in practice, as it provides a good fit to market data while remaining stable even when data is limited, especially relevant for our dataset. Moreover, its reduced number of parameters ensures a faster and more robust calibration. Additionally, its structure allows for smooth extrapolation across maturities, making it especially useful in real-world applications where data may be sparse or irregular.

Our results showed that two of these parameters, a and p, can be at least partially predicted using the historical path of the underlying asset. The model performed especially well for the parameter a, while predictions for p improved significantly after filtering the dataset to retain only ATM options.

Overall, our findings support the idea that implied volatility is, not only influenced by current market conditions, but also by the long-term evolution of the underlying index, adding a dimension to predictive techniques for implied volatility surfaces and further highlighting the usefulness of simple path-dependent models, especially for the ATM region. Future work could focus on improving the modeling of the remaining parameters of the Parsimonious SSVI and better capturing market regime changes.

References

[1] Hervé Andrès, Alexandre Boumezoued, and Benjamin Jourdain. "Implied volatility (also) is path-dependent". In: *arXiv preprint arXiv:2312.15950* (2023).

[2] SciPy Developers. *SciPy: Scientific Library for Python*. `https://docs.scipy.org/doc/scipy/`. Accessed: 2025-04-17. 2024.

[3] SciPy Developers. *SLSQP Method — SciPy Optimize Documentation*. Accessed: 2025-04-17. 2024. URL: `https://docs.scipy.org/doc/scipy/reference/generated/scipy.optimize.minimize.html#slsqp`.

[4] Jim Gatheral and Antoine Jacquier. "Arbitrage-free SVI volatility surfaces". In: *Quantitative Finance* 14.1 (2014), pp. 59–71.

[5] Julien Guyon and Jordan Lekeufack. *Volatility Is (Mostly) Path-Dependent*. 2023. URL: `https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4174589`.

[6] Dieter Kraft. "A software package for sequential quadratic programming". In: *Forschungsbericht-Deutsche Forschungs- und Versuchsanstalt fur Luft- und Raumfahrt* (1988).

[7] Jordan Lekeufack Sopze. *Volatility is (Mostly) Path Dependent*. `https://github.com/Jordylek/VolatilityIsMostlyPathDependent`. Accessed: 2025-04-17. 2024.