Therefore, we can greatly increase the number of training scenarios, since now the hedging scenarios can have arbitrary combinations of T and K even if the combinations of T and K are not directly observed in the market.

6.1.1 No-Arbitrage Price Surface

In this section, we describe how to calibrate an arbitrage-free price surface on each business day. We want to create a price surface that is **arbitrage-free**. In addition, we want to be able to **extrapolate** for in-money options and out-of-money options. In the following subsection, we will discuss the steps of calibration of the arbitrage-free surface with.

Following [80], we say that an implied volatility surface is free of *static arbitrage* if and only if it is free of *call spread arbitrage*, *calendar arbitrage*, and *butterfly arbitrage*, which we describe below. We assume we have a collection of European call option prices $\{C(T,K)\}_{T,K}$ for a range of strikes, K, and expiries, T, from which we associate its Black-Scholes implied volatility surface $\{\sigma^{imp}(T,K)\}_{T,K}$. We also suppose that interest rates are deterministic with the expressions $D(t,T) = e^{-r(T-t)}$ for discount factors, t be trading date and T be the maturity date.

1. <u>Butterfly Arbitrage</u>: Given a collection of call option prices $\{C(T,K)\}_{T,K}$, using Dupire's method or equivalent approaches [64] one can (at least formally) find expressions for implied probability densities $\{p(\cdot;T,S_0)\}_T$ such that

$$C(T,K) = D(t,T)\int_{(0,\infty)} (S-K)_+ p(S;T,S_t)dS.$$

We say that the surface $\{\sigma^{imp}(T,K)\}_{T,K}$, where $\sigma^{imp}(T,K)$ is the black-scholes implied volatility, is free of Butterfly Arbitrage if its implied probability densities $\{p(\cdot;T,S_0)\}_T$ are valid densities, i.e., $p(S;T,S_t) \geq 0$ for all $S > 0$ and $\int_0^\infty p(S;T,S_t)dS = 1$. Additionally, the condition $p(S;T,S_t) \geq 0$ for all $S > 0$ is equivalent to require $\frac{\partial^2 C(T,K)}{\partial K^2} > 0$ for all $K > 0$ since Breeden and Litzenberger [23] show that:

$$\left.\frac{\partial^2 C(T,K)}{\partial K^2}\right|_{K=x} = D(t,T)p(x;T,S_t)$$

2. <u>Calendar Arbitrage</u>: Given a surface $\sigma^{imp}(T,K)$, we consider the corresponding total variance (TV) surface defined by

$$w(\tau,k) = \sigma(T,K)^2\tau$$

where $\tau = T-t$ and k is parameterized by log-moneyness, i.e. $k := \log(K/F(t;T))$ and $F(t;T) = S_t e^{(r-q)(T-t)}$
is the at-the-money forward (ATMF) price for S_t. We say that the surface $\{\sigma^{imp}(T,K)\}_{T,K}$ is free of Calendar Arbitrage

- if $\frac{\partial w(\tau,k)}{\partial T} \geq 0$ for all $k \in \mathbb{R}, T > 0$ for continuous time data
- if $w(\tau_i, k) \leq w(\tau_{i+1}, k)$ for all $k \in \mathbb{R}, T_i \leq T_{i+1}$ for discrete time data

Furthermore, given a grid of strikes: $0 = K_0 < K_1 < \cdots < K_{max}$, and a grid of expiries $t = T_0 < T_1 < \cdots < T_{max}$. Let $C(T,K)$ be the call option price of the expiry T and strike K returned by a model. The corresponding discrete criteria [31] for a grid of option prices to be free of arbitrage are set as the following:

1. No butterfly spread arbitrage condition:

$$C(T_j, K_{i-1}) - C(T_j, K_i) > \frac{K_i - K_{i-1}}{K_{i+1} - K_i} \left(C(T_j, K_i) - C(T_j, K_{i+1}) \right) \tag{6.1.1}$$

2. No calendar spread arbitrage:

$$C(T_j, K_i) \geq C(T_{j-1}, K_i) \tag{6.1.2}$$

Volatility Interpolation and Extrapolation Over Strikes

In this thesis, we will use SABR model to obtain the option price for an arbitrary strike K unobserved in the market. On each business day t, we have observed option quotes for a set of market observed expires. We denote the market available expires as.

$$\mathbf{T}_t^{mkt} = \{T_0^t < T_1^t < \cdots < T_{max}^t\}$$

The first expiry is set to be t: $T_0^t = t$. Note that for the first expiry $T_0^t = t$, the market option value at t is just the option payoff at t which is $\max\{S_t - K, 0\}$ (call) or $\max\{K - S_t, 0\}$ (put).

Since in this subsection, we are discussing how to obtain a arbitrage-free surface on a fixed date t, for the notational simplicity, we will omit the superscript and subscript t in this subsection, we denote the grid of market observed expires on t to be

$$\{T_0 < T_1 < \cdots < T_{max}\}$$

We firstly use SABR model to obtain option price for a predetermined grid of strike $0 = K_0 < K_1 < \cdots < K_{max}$. More specifically, for a fixed expiry T_i, let us define

$$K_{T_i}^{mkt} = \{K_1^{mkt}, \ldots, K_{N_K}^{mkt}\}$$

to be the strikes for which market option price is available for T_i. We solve

$$\min_{\alpha,\nu,\rho} \sum_{j=1}^{N_K} \left(V_{SABR}(S, t, T_i, K_j^{mkt}, r, q; \alpha, \beta, \nu, \rho) - V^{mkt}(t, T_i, K_j^{mkt}) \right)^2$$

where $V_{SABR}(S,t,T,K,r,q;\alpha,\beta,\nu,\rho)$ is option pricing function described in section 2.2.3. Since we are dealing with S&P500 index option, the log-normal assumption is more appropriate, we fix $\beta = 1$ in the calibration process. We choose to calibrate a different model for each expiry. A different set of parameters is specified for each expiry, describing an instantaneous process. The SABR models for different expiries are completely independent from each other. We choose this approach because the single implied volatility surface calibrated for all expiries and strikes is unlikely to fit the actual surface very well and calibrating a single surface is harder and more time-consuming.

Note that we have three parameters to be calibrated so we need to observe at least 3 data points from market to successfully build a SABR model. Therefore, the number of strikes N_K must be larger than or equal to 3. [2]

Although SABR model is computationally efficient and can match the market volatility smile well, it is not arbitrage-free. The formula (2.2.11) is an approximation, obtained from an asymptotic series expansion. Its accuracy degrades if the option strikes move away from the option at-the-money (ATM) strike. As a result, the implied probability density function:

$$\frac{1}{D(t,T)} \left. \frac{\partial^2 C_{SABR}(T,K)}{\partial K^2} \right|_{K=x} = p(x;T,S_t)$$

where $C_{SABR}(T,K)$ is the SABR price of a call option at strike K and expiry T, may become negative at very low or very high strikes. Therefore, we may have butterfly spread arbitrage in SABR prices returned by the calibrated models. In addition, for each expiry, a separate set of SABR parameters is calibrated so that calender spread arbitrage can also exist. However, the existence of calendar arbitrage will be rare since the market option prices rarely contains calendar arbitrage.

Therefore, if we found that the grid of prices returned by the SABR model has failed any of the **discrete** arbitrage conditions, we will introduce some adjustments. To fix the butterfly spread arbitrage, we implements an risk-neutral adjustment [26]. This adjustment substitutes the implied distribution tails by those of certain log-normal distributions. Interested reader can refer to [26] for more details. Here we just briefly discuss the process of the adjustment.

First, we introduce lower and upper strike limits, K_L and K_U within which the implicit probability density function (p.d.f) $p(x;T,S_t)$ is assumed to be valid. The lower and upper strike limit can be the maximum and minimum strike K for which the discrete no butterfly spread arbitrage condition (6.1.2) holds . Then we assume the adjusted p.d.f is of the form:

$$\hat{g}(x) = \begin{cases} \lambda_L \, q(x;\mu_L,\sigma_L), & \text{if } 0 < x < K_L \\ p(x;T,S_t), & \text{if } K_L \le x \le K_U \\ \lambda_U \, q(x;\mu_U,\sigma_U), & \text{if } x > K_U \end{cases}$$

[2] The $V^{mkt}(t,T_i,K_j^{mkt})$ is the mid-price of market observed best bid and best ask prices. We use market quotes in calibration and even if the trading volume is zero, we will still include them in the calibration.

where $q(x;\mu,\sigma)$ is the p.d.f of a log-normal distribution:

$$q(x;\mu,\sigma) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{(\log(x)-\mu)^2}{2\sigma^2}}.$$

and $p(x;T,S_t)$ is the implied probability density from the calibrated SABR model. We require the following condition to be satisfied:

1. Intergrability constraint

$$\int_0^{K_L} \lambda_L q(x;\mu_L,\sigma_L)dx + \int_{K_L}^{K_U} p(x;T,S_t)dx + \int_{K_U}^{\infty} \lambda_U\, q(x;\mu_U,\sigma_U)dx = 1 \tag{6.1.3}$$

2. Martingale constraint

$$\int_0^{K_L} x\lambda_L q(x;\mu_L,\sigma_L)dx + \int_{K_L}^{K_U} xp(x;T,S_t)dx + \int_{K_U}^{\infty} x\lambda_U\, q(x;\mu_U,\sigma_U)dx = F(t,T) \tag{6.1.4}$$

We solve the equations (6.1.3) and (6.1.4) for the unknown parameters $\{\mu_L,\mu_U,\sigma_L,\sigma_U,\lambda_L,\lambda_U\}$.

Let:

- $C_{BS}(T,K;\sigma)$ and $P_{BS}(T,K;\sigma)$ be the call and put option prices from Black-Scholes model with the Black-Scholes volatility σ.
- $DC_{BS}(T,K;\sigma)$ and $DP_{BS}(T,K;\sigma)$ be the digital call and put option prices from Black-Scholes model with the Black-Scholes volatility σ where a digital call pays one dollar if the underlying price exceeds the strike and a digital put pays the same amount if the underlying is below the strike.
- σ_B be implied Black's volatility given by the SABR apprximation formula (2.2.11) for $K \in [K_L, K_U]$. For simplicity, we write $\sigma_B(K)$ to denote the implied Black's volatility given by the SABR formula (2.2.11) for a strike K.

Additionally, observing that the Black-Scholes [19] model implies that the prices of the underlying asset S_T at the maturity T are log-normal distribued:

$$\ln(S_T) \sim \mathcal{N}(\ln(S_t) + (r-q-\frac{\sigma^2}{2})(T-t), \sigma^2(T-t))$$

we set $\mu_L = \ln(S_t) + (r-q-\frac{\sigma_L^2}{2})(T-t)$ and $\mu_U = \ln(S_t) + (r-q-\frac{\sigma_U^2}{2})(T-t)$ such that we have only four parameters to be solved: $\{\sigma_L, \sigma_U, \lambda_L, \lambda_U\}$. Furthermore, by such setting, we will have:

$$
\begin{aligned}
D(t,T)\int_0^{K_L} \lambda_L q(x;\mu_L,\sigma_L)dx &= \lambda_L DP_{BS}(T,K_L;\sigma_L) \\
D(t,T)\int_{K_U}^{\infty} \lambda_U\, q(x;\mu_U,\sigma_U)dx &= \lambda_U DC_{BS}(T,K_U;\sigma_U) \\
D(t,T)\int_0^{K_L} (K_L - x)\lambda_L q(x;\mu_L,\sigma_L)dx &= \lambda_L P_{BS}(T,K_L;\sigma_L) \\
D(t,T)\int_{K_U}^{\infty} (x - K_U)\lambda_U\, q(x;\mu_U,\sigma_U)dx &= \lambda_U C_{BS}(T,K_U;\sigma_U)
\end{aligned}
\tag{6.1.5}
$$

Following equation (6.1.5), we observe that the two constraints (6.1.3) and (6.1.4) can be written as:

$$
\begin{aligned}
\frac{1}{D(t,T)} =& \lambda_L DP_{BS}(T,K_L;\sigma_L) + \\
& \frac{1}{D(t,T)} - DP_{BS}(T,K_L;\sigma_B(K_L)) - DC_{BS}(T,K_U;\sigma_B(K_U)) + \\
& \lambda_U DC_{BS}(T,K_U;\sigma_U)
\end{aligned}
\tag{6.1.6}
$$

$$
\begin{aligned}
F(t,T) =& \frac{1}{D(t,T)}\lambda_L[K_L DP_{BS}(T,K_L;\sigma_L) - P_{BS}(T,K_L;\sigma_L)] + \\
& F(t,T) - \frac{1}{D(t,T)}[K_L DP_{BS}(T,K_L;\sigma_B(K_L)) - P_{BS}(T,K_L;\sigma_B(K_L))] - \\
& \frac{1}{D(t,T)}[K_U DC_{BS}(T,K_U;\sigma_B(K_U)) + C_{BS}(T,K_U;\sigma_B(K_U))] + \\
& \frac{1}{D(t,T)}\lambda_U[K_U DC_{BS}(T,K_U;\sigma_U) + C_{BS}(T,K_U;\sigma_U)]
\end{aligned}
\tag{6.1.7}
$$

As one can easily see, if we have:

$$
\begin{aligned}
\lambda_L DP_{BS}(T,K_L;\sigma_L) &= DP_{BS}(T,K_L;\sigma_B(K_L)) \\
\lambda_U DC_{BS}(T,K_L;\sigma_U) &= DC_{BS}(T,K_U;\sigma_B(K_U)) \\
\lambda_L P_{BS}(T,K_L;\sigma_L) &= P_{BS}(T,K_L;\sigma_B(K_L)) \\
\lambda_U C_{BS}(T,K_L;\sigma_U) &= C_{BS}(T,K_U;\sigma_B(K_U))
\end{aligned}
\tag{6.1.8}
$$

we can show the intergrability constraint (6.1.6) and martingale constraint (6.1.7) are satisified.

Therefore, we can solve equation (6.1.8) easily with a non-linear least square solver for the parameter: $\{\sigma_L, \sigma_U, \lambda_L, \lambda_U\}$. After the calibration, we will simply using the adjusted p.d.f for option pricing:

$$
C_{SABR}(T,K) = D(t,T)\int_{-\infty}^{\infty} (x-K)_+ \hat{g}(x)dx \tag{6.1.9}
$$

We can easily verify that (6.1.9) can be written as:

$$C_{SABR}(T,K) = \begin{cases} \lambda_L C_{BS}(T,K;\sigma_L) & \text{if } 0 < K < K_L \\ C_{BS}(T,K;\sigma_B(K)) & \text{if } K_L \le K \le K_U \\ \lambda_U C_{BS}(T,K;\sigma_U) & \text{if } K > K_U \end{cases}$$

The similar logic will be used to correct the put option prices if any arbitrage exists.

For calendar arbitrage, we will shift the price by the the following procedure to remove it.

$$shift_{t_j} = -min(min_K(C_{SABR}(T_j,K) - C_{SABR}(T_{j-1},K)),0).$$

$$C_{SABR}(T_j,K) \leftarrow C_{SABR}(T_j,K) + shift_{t_j}$$

Note the shift will be zero if no calendar arbitrage is observed between T_j and T_{j-1}.

Alternative Methods to SABR Calibration

There are many other alternative methods to calibrate the arbitrage-free surface as SABR model. One of the most successful volatility surface models is the Stochastic Inspired Volatility model (SVI) [78] used internally at Merrill Lynch and publicly disclosed by Jim Gatheral in 2004. SVI is a simple five-parameters formula. Numerous applied research teams have attempted to extend SVI to a whole surface model. This has been achieved in 2012 by Gatheral and Jacquier with the Surface SVI (SSVI) [80]. SSVI is parameterized in a way that a SSVI slice at a given maturity T is a SVI slice with only 3 parameters. This restriction allows to get explicit sufficient conditions for absence of arbitrage, while allowing enough flexibility for calibration. The SSVI model is recently extended by [96] and [50]. If the input market data used for calibration contains arbitrage, the calibrated surface from SSVI [80] or its extension [96, 50] can typically be viewed as the surface that is as close as possible to the original market data, while staying arbitrage-free. In this section we briefly review a recent variant dd-eSSVI method [50] as an example.

Given volatility surface market data as a discrete set of points $\{\sigma^{imp}(T_j,K_i)\}_{i,j}$, a natural question is whether static arbitrage exists. In most cases, one must

1. Fit a parameterization to the points $\{\sigma^{imp}(T,K)\}_{T,k}$, typically by time-slice as what we did with SABR model, obtaining a parameterization $\{\widetilde{\sigma}(T_i,K)\}_{T_i}$ for each T_i separately.
2. Compute $\{\widetilde{w}(\tau_i,k)\}_{i=1}^{N}$ on a fine grid of k using the expression $\widetilde{w}(\tau_i,k) := \widetilde{\sigma}(T_i,K)^2 T_i$
3. Check for the properties listed above defining condition 1 butterfly arbitrage and condition 2 calendar arbitrage or the corresponding discrete criteria as in [31].

To obtain such mappings, one typically fits a stochastic volatility model like SABR or Heston to slices of $\{\sigma^{mkt}(K,T)\}_{K,T}$ or considers a generalized model such as Stochastic Volatility Inspired (SVI) parameterizations.

SVI Parameterization

Following the work of [50], we introduce the following Stochastic Volatility Inspired (SVI) parameterization for a surface's Total Variance (TV), $w(\tau,k)$ as

$$w(\tau,k) = \frac{1}{2}\left(\hat{\theta}_\tau + \hat{\rho}_\tau \hat{\psi}_\tau k + \sqrt{\left(\hat{\psi}_\tau k + \hat{\rho}_\tau \hat{\theta}_\tau\right)^2 + \left(1-\hat{\rho}_\tau^2\right)\hat{\theta}_\tau^2}\right) \tag{6.1.10}$$

In this parameterization we have that

- $\hat{\theta}_\tau$ is the ATM Forward TV, read directly from the surface

  $$\hat{\theta}_\tau = w(\tau,0) = \sigma(T,K_{ATMF})^2 \cdot \tau$$

  where $K_{ATMF} = F(t,T)$
- $\hat{\rho}_\tau$ controls the slope of the skew
- $\hat{\psi}_\tau$ controls the curvature.

An important feature of this parameterization is that it provides easy to impose sufficient conditions on the parameters so that there is no butterfly arbitrage for a given slice, and no calendar arbitrage between two time slices. These conditions are listed below

Arbitrage Conditions

- Butterfly Arbitrage: The surface is free of Butterfly Arbitrage if for each time-slice

  $$0 < \hat{\psi} < \min\left(\sqrt{\frac{4\hat{\theta}}{1+|\hat{\rho}|}}, \frac{4}{1+|\hat{\rho}|}\right) \tag{6.1.11}$$

  for $\hat{\theta} > 0, \hat{\rho} \in (-1,1)$.
- Calendar Arbitrage: The surface is free of Calendar Arbitrage under the following set up: Given $\tau_1 \leq \tau_2$, we let $\hat{\theta}_1 := \hat{\theta}_{\tau_1}, \hat{\theta}_2 := \hat{\theta}_{\tau_2}$, $\hat{\psi}_1 := \hat{\psi}_{\tau_1}, \hat{\psi}_2 := \hat{\psi}_{\tau_2}$. To be free of calendar arbitrage it is necessary that

  $$\hat{\theta}_{1,2} \geq 1 \text{ and } \hat{\psi}_2 > \max\left(\frac{\hat{\psi}_1 + \hat{\rho}_1\hat{\psi}_1}{1+\hat{\rho}_2}, \frac{\hat{\psi}_1 - \hat{\psi}_1\hat{\rho}_1}{1-\hat{\rho}_2}\right) \tag{6.1.12}$$

  If the above necessary condition holds, it is further sufficient that

  $$\hat{\psi}_{1,2} \leq \hat{\theta}_{1,2} \text{ or } (\hat{\rho}_1 - \hat{\psi}_{1,2}\hat{\rho}_2)^2\hat{\theta}_{1,2} \leq (\hat{\theta}_{1,2}-1)(\hat{\psi}^2 - \hat{\theta}_{1,2}). \tag{6.1.13}$$

  in order to preclude calendar arbitrage.