

SABR Model for the Implied Volatility Smile

Published by **BSIC** on 4 November 2018

The SABR model is like the Vega/Vanna Volga Approach, in that it is a method of interpolating the implied volatility surface. Given the dynamics of the forward rate, the stochastic instantaneous volatility, and the Black model, we get an algebraic expression that the Black Implied Volatility must satisfy.

Given traded and liquid options, we fit the SABR model on the observed smile and estimate the parameters. Using these parameters, we can estimate Black Implied Vol to price exotics and vanilla swaptions at various points on the vol grid. The underlying assumption is: let the market tell us what to do. We take the opinion of the market on liquid points on the vol grid to estimate parameters and extrapolate/interpolate them to other illiquid/non-traded points in order to price options.

The SABR model

The SABR model assumes that the forward rate and the instantaneous volatility are driven by two correlated Brownian motions:

$$\begin{aligned} df_t &= \alpha_t f_t^\beta dW_t^1 \\ d\alpha_t &= v \alpha_t dW_t^2 \\ E[dW_t^1 dW_t^2] &= \rho dt \end{aligned}$$

The expression that the implied volatility must satisfy is ¹



$$(fK)^{(1-\beta)/2} \left[1 + \frac{(1-\beta)^{\alpha}}{24} \ln^2 \frac{f}{K} + \frac{(1-\beta)^{\alpha}}{1920} \ln^4 \frac{f}{K} \right] \\ \times \frac{z}{\chi(z)}$$

$$\begin{aligned} z &= \frac{v}{\alpha} (fK)^{(1-\beta)/2} \ln \frac{f}{K} \\ \chi(z) &= \ln \left[\frac{\sqrt{1 - 2\rho z + z^2} + z - \rho}{1 - \rho} \right]. \end{aligned}$$

When $f = K$ (for ATM options), the above formula for implied volatility simplifies to:

$$\sigma_{ATM} = \sigma_B(f, f) = \frac{\alpha \left\{ 1 + \left[\frac{(1-\beta)^2}{24} \frac{\alpha^2}{f^{2-2\beta}} + \frac{1}{4} \frac{\rho \beta v \alpha}{f^{1-\beta}} + \frac{2-3\rho^2}{24} v^2 \right] T \right\}}{f^{1-\beta}}$$

where

- α is the instantaneous vol;
- v is the vol of vol;
- ρ is the correlation between the Brownian motions driving the forward rate and the instantaneous vol;
- β is the CEV component for forward rate (determines shape of forward rates, leverage effect and backbone of ATM vol).

The local volatility function, which includes the stochastic instantaneous vol and the deterministic function of F , expressed as % of forward rate is given by:

$$\hat{\alpha} \hat{F}^\beta / \hat{F}^1 = \hat{\alpha} \hat{F}^{\beta-1}$$

β is one of the key parameters and affect many fundamental characteristics of the model. β is the exponent for the forward rate or the CEV exponent and effects the distribution

- $\beta = 1$ – stochastic lognormal rates

As the CEV component, β helps capture leverage effect without changing the underlying instantaneous vol (α). So, if $0 < \beta < 1$ (as in SABR), we have leverage effect –volatility falls as F increases.

The dynamics of the Implied Volatility formula

For strikes not too far from the forward rate, the Implied Volatility formula can be approximated as

$$\begin{aligned}\sigma_B(K, f) = & \frac{\alpha}{f^{1-\beta}} \left\{ 1 - \frac{1}{2} (1 - \beta - \rho\lambda) \log K/f \right. \\ & \left. + \frac{1}{12} [(1 - \beta)^2 + (2 - 3\rho^2)\lambda^2] \log^2 K/f + \dots \right.\end{aligned}$$

where λ measures the strength of vol-of-vol against the local volatility at the current forward rate

$$\lambda = \frac{\nu}{\alpha} f^{1-\beta}$$

The dynamics of $\sigma_B(K, f)$ can be broken down into 3 parts – backbone, smile(convexity) and skew.

1- The backbone

The backbone $\frac{\alpha}{f^{1-\beta}}$ is the change in ATM vol for a change in the forward rate. The backbone is almost entirely determined by the exponent β : with $\beta = 0$ (a stochastic vol Gaussian model) we get a steeply downward sloping backbone, and with $\beta = 1$ (a stochastic vol lognormal model) we get a nearly flat backbone – as in the Black and Scholes model. To plot the backbone for given values of α , β , ρ and ν , we change the value of f and compute the ATM vol assuming that $K=f$ (ie, using the ATM vol formula).

2- The skew

The skew is given by $-\frac{1}{2}(1 - \beta - \rho\lambda)\log\left(\frac{K}{f}\right)$ and is the sum of two components:

- the **Beta Skew** $-\frac{1}{2}(1 - \beta)\log\left(\frac{K}{f}\right)$, which makes the implied vol downward sloping in terms of K (because $0 < \beta < 1$). It arises because of the leverage effect built into the local volatility (assuming constant instantaneous vol) – ie, because the local volatility $\alpha/f^{1-\beta}$ is a decreasing function of the forward price;

The volatility α would decrease (increase) when the forward f increases (decreases). It thus seems unsurprising that a negative correlation ρ causes a downward sloping vanna skew (lower vol at higher strikes).

Interestingly, the slope of the backbone will be twice as steep as the slope of the smile due to the β -component of the skew. Basically, according to the backbone, the ATM vol falls as f increases. The skew is also affected by the backbone multiplier – so as the f increases, the beta component of skew makes the implied vol at a K rise, but ATM vol falls due to backbone: thus the impact on skew is dampened.

3- The convexity

The convexity term responsible for the smile is

$$+ \frac{1}{12} [(1 - \beta)^2 + (2 - 3\rho^2) \lambda^2] \log^2 K/f.$$

and can also be decomposed into two terms:

- the smile induced by the **Skew effect** $\frac{1}{12} (1 - \beta)^2 \log^2 \left(\frac{K}{f}\right)$. It is essentially a second order effect of the skew term for bigger moves from $K = f$;
- the smile induced by the **Volga effect** $\frac{1}{12} (2 - 3\rho^2) \lambda^2 \log^2 \left(\frac{K}{f}\right)$

$$(2 - 3\rho^2) \lambda^2] \log^2 K/f.$$

Intuitively, the smile arises because unusually large movements of the forward f happen more often when the volatility α increases, and less often when α decreases, so strikes K far from f represent on average, high volatility environments.

How to fit the Market Data

1- Beta

β has little effect on the shape of the surface/smile (as a snapshot at a given f and time to expiry), but:

Bartlett's delta – to reduce the sensitivity of delta to the choice of β (more on this later).

To elaborate on the first point above, market smiles can be fit equally well with any specific value of β . Because The exponent β and correlation ρ affect the volatility smile in similar ways – (they both cause a downward sloping skew in $\sigma_B(K, f)$ as the strike K varies), from a single market snapshot of $\sigma_B(K, f)$ as a function of K at a given f , it is difficult to distinguish between the two parameters.

For instance, in the figure below we fit the SABR parameters α, ρ, v with $\beta = 0$ and, then, re-fit the parameters α, ρ, v with $\beta = 1$. Note that there is no substantial difference in the quality of the fits, despite the presence of market noise.

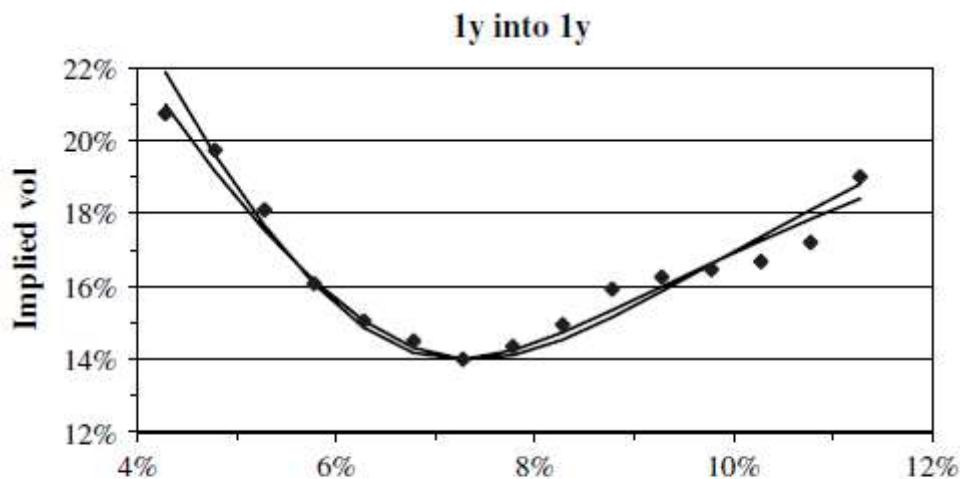


Fig. 3.3 Implied volatilities as a function of strike. Shown are the curves obtained by fitting the SABR model with exponent $\beta = 0$ and with $\beta = 1$ to the 1y into 1y swaption vol observed on 4/28/00. As usual, both fits are equally good. Data courtesy of Cantor-Fitzgerald.

Because β has little effect on the shape of the surface, beta is usually set using a regression or according to our belief of the distribution of rates, and then we estimate the other parameters. Hence, we generally don't fit smiles to estimate beta. β can be determined from:

- historical observations of the backbone. Taking the log of the main SABR ATM vol formula, we get

$$\log \sigma_B(f, f) = \log \alpha - (1 - \beta) \log f + \log \left\{ 1 + \left[\frac{(1 - \beta)^2}{24} \frac{\alpha^2}{f^{2-2\beta}} + \frac{1}{4} \frac{\rho \beta \alpha v}{f^{(1-\beta)}} + \frac{2 - 3\rho^2}{24} v^2 \right] t_{ex} + \dots \right\}$$

- “aesthetic considerations”, based on our belief of the distribution of f :
 - $\beta = 1$ (stochastic lognormal) – proponents say log normal models are “more natural” or believe that the horizontal backbone best represents their market. These proponents often include desks trading foreign exchange options;
 - $\beta = 0$ (stochastic normal, with negative rates) – Proponents usually believe that a normal model, with its symmetric break-even points, is a more effective tool for managing risks, and would claim that $\beta = 0$ is essential for trading markets like Yen interest rates, where the forwards f can be negative or near zero;
 - $\beta = 1/2$ (stochastic CIR) models – usually US interest rate desks that have developed trust in CIR models.

2- Other parameters

With β given, fitting the SABR model is a straightforward procedure. The three parameters α , ρ , and v have different effects on the curve:

- α controls the overall height of the curve;
- ρ controls the curve’s skew;
- v controls how much smile the curve exhibits.

Because of the widely separated roles these parameters play, the fitted parameter values tend to be very stable, even in the presence of large amounts of market noise. It is usually more convenient to use the at-the-money volatility σ_{ATM} , β , ρ , and v as the SABR parameters instead of the original parameters α , β , ρ , v . In the generic volatility formula of SABR we express all the α terms as a function of σ_{ATM} (basically we manipulate the equation such that α is removed). The parameter α is then found whenever needed by inverting the main ATM Implied Vol formula on the fly (given below). This inversion is numerically easy since the higher order terms are small.

$$\sigma_{ATM} = \sigma_B(f, f) = \frac{\alpha}{f^{(1-\beta)}} \left\{ 1 + \left[\frac{(1-\beta)^2}{24} \frac{\alpha^2}{f^{2-2\beta}} + \frac{1}{4} \frac{\rho \beta \alpha v}{f^{(1-\beta)}} + \frac{2-3\rho^2}{24} v^2 \right] t_{ex} + \dots \right\}$$

- this parameterization, fitting the SABR model requires fitting ρ and v to the implied volatility
- , with σ_{ATM} and β given. In many markets, the ATM volatilities need to be updated frequently,

Some observation from fitting SABR on US swaptions and rates options market

The smile and skew depend heavily on the time-to exercise. Smile dominates in shorter maturities, as the vol of vol ν is very high for short dated options, and decreases as the time-to-exercise increases. This matches general experience: in most markets there is a strong smile for short-dated options which relaxes as the time-to-expiry increases; consequently the volatility of volatility ν is large for short dated options and smaller for long-dated options, regardless of the particular underlying.

Experience with correlations is less clear: in some markets a nearly flat skew for short maturity options develops into a strongly downward sloping skew for longer maturities. In other markets there is a strong downward skew for all option maturities, and in still other markets the skew is close to zero for all maturities.

For swaption, there is little dependence of the market skew and smile on the length of the underlying swap: both ν and ρ are fairly constant across tails.

The predicted dynamics of the smile matches market experience – whenever the forward price f changes, the implied volatility curve shifts in the same direction and by the same amount as the price f – this overcomes a massive problem with local vol models where the smile shifts in the opposite direction as f .

If $\beta < 1$, the “backbone” is downward sloping, so the shift in the implied volatility curve is not purely horizontal – the curve shifts up and down as the at-the-money point traverses the backbone.

Managing Smile Risk and the modified greeks

A big advantage of risk management with SABR is that greeks are consistent across any one smile (at a given maturity) and can be added. Since the SABR model is a single self-consistent model for all strikes K , the risks calculated at one strike are consistent with the risks calculated at other strikes. Therefore, the risks of all the options on the same asset and with the same maturity can be added together, and only the residual risk needs to be hedged. On the other hand, in the Black formula each strike has a different constant volatility, which essentially means we are using a different model at each K , which makes the risks not directly comparable across strikes.

According to the SABR model, the value of a call is $V_{call} = BS(f, K, \sigma_B(K, f), t_{ex})$ where the volatility $\sigma_B(K, f) = \sigma_B(K, f; \alpha, \beta, \rho, \nu)$ is the SABR implied volatility. In other words, we can

Delta

1- Delta in original SABR

If we use the parameterization α, β, ρ, ν , so that $\sigma_B(K, f) = \sigma_B(K, f; \alpha, \beta, \rho, \nu)$,

$$\Delta \equiv \frac{\partial V_{call}}{\partial f} = \frac{\partial BS}{\partial f} + \frac{\partial BS}{\partial \sigma_B} \frac{\partial \sigma_B(K, f; \alpha, \beta, \rho, \nu)}{\partial f}.$$

The first term is the regular Black delta.

The second term is the SABR model's correction to the risk – Black vega*times the predicted change in the implied volatility σ_B caused by the change in the forward f . This change consists of –

- The option moving and down the smile – ie, caused by a sideways movement of the volatility curve in the same direction (and by the same amount) as the change in the forward price f .
- If $\beta < 1$ the volatility curve rises and falls as the at-the-money point traverses up and down the backbone.

Note – if we use the parameterization $\sigma_{ATM}, \beta, \rho, \nu$, delta is written as –

$$\Delta \equiv \frac{\partial BS}{\partial f} + \frac{\partial BS}{\partial \sigma_B} \left\{ \frac{\partial \sigma_B(K, f; \alpha, \beta, \rho, \nu)}{\partial f} + \frac{\partial \sigma_B(K, f; \alpha, \beta, \rho, \nu)}{\partial \alpha} \frac{\partial \alpha(\sigma_{ATM}, f)}{\partial f} \right\}.$$

Where α is a function of σ_{ATM} and f defined by the SABR ATM vol formula. The delta risk is now the risk to changes in f with σ_{ATM} held fixed. The last term is just the change in α needed to keep σ_{ATM} constant while f changes. Clearly this last term must just cancel out the vertical component of the backbone, leaving only the sideways movement of the implied volatility curve. Note that this term is zero for $\beta = 1$.

Theoretically, one should use the delta from the first parameterization to risk-manage books. In many markets, however, it may take several days for volatilities σ_B to change following significant changes in the forward price f . In these markets, using the second delta is a much more effective hedge.

For suppose one used from equation the first delta – when the volatility σ_{ATM} did not immediately following a change in f , one would be forced to re-mark α to compensate, and this re-marking would change the hedges. As σ_{ATM} equilibrated over the next few days, one would mark α back to its

Let us use the first delta formula for the rest of this section:

$$\Delta \equiv \frac{\partial V_{call}}{\partial f} = \frac{\partial BS}{\partial f} + \frac{\partial BS}{\partial \sigma_B} \frac{\partial \sigma_B(K, f; \alpha, \beta, \rho, v)}{\partial f}.$$

The adjustment term to Black Delta here only accounts for changes caused by the local vol function – that is, it accounts for systematic changes in implied vol as f changes (implicitly assuming that alpha remains constant). However, due to a non-zero correlation between alpha and f , the alpha changes too as f changes and hence the change in implied vol is both a function of implied volatility changing deterministically as f changes (as measured by local vol) as well as the underlying stochastic vol (alpha) changing as f changes. This correction is made by Barlett.

2- Updated Bartlett Delta to account for changes in instantaneous vol caused by f

In the original SABR model, the delta hedge was calculated by shifting the current value of the underlying while keeping the current value of α fixed:

$$f \rightarrow f + \Delta f, \alpha \rightarrow \alpha.$$

However, since α and f are correlated, whenever f changes, on average α changes as well. A delta scenario which is more realistic than the above scenario is

$$f \rightarrow f + \Delta f, \alpha \rightarrow \alpha + \delta_f \alpha$$

The average change in α due to a change in the forward is given by

$$\delta_f \alpha = \frac{\rho v}{f^\beta} \Delta f$$

Hence, the new delta risk is given by

$$\Delta = \frac{\partial B}{\partial f} + \frac{\partial B}{\partial \sigma} \left(\frac{\partial \sigma}{\partial f} + \frac{\partial \sigma}{\partial \alpha} \frac{\rho v}{f^\beta} \right)$$

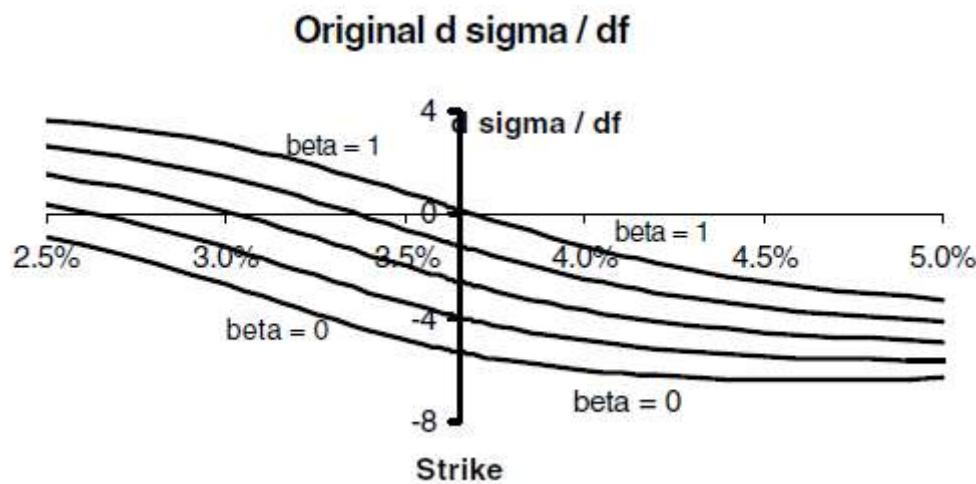
This incorporates the average change in instantaneous volatility caused by change in the underlying. The new term in the updated delta risk is just $\rho v/f^\beta$ times the classic SABR vega risk.

If the (classic SABR) vega risk and (classic SABR) delta risk are both hedged, then the Bartlett delta is also hedged. However, if we only hedge the classic SABR delta, this term new won't be hedged.

major impact on the delta.

Remember that the original SABR delta only incorporated the volatility risk due to f changing through the local volatility – systematic change in the implied vol curve due to changes in the forward f expressed as $\partial\sigma/\partial f$.

The figure below shows the $\partial\sigma/\partial f$ term from the original SABR article.



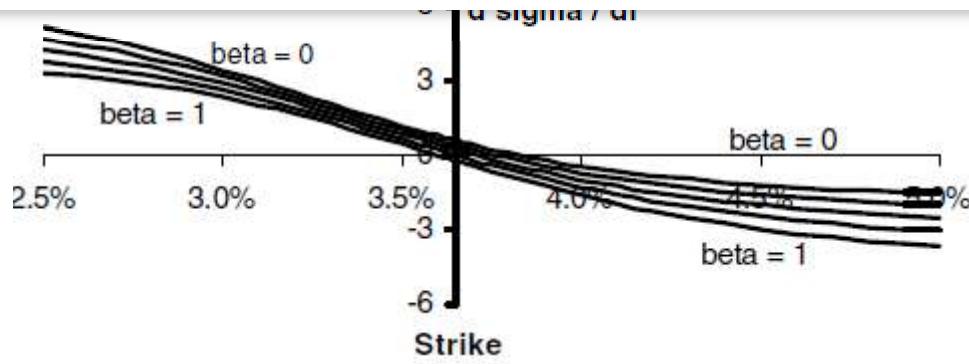
This term changes substantially with the value of beta, causing the Classic SABR delta's sensitivity to beta.

Now, the Bartlett adjusted delta accounts for the average change in α that occurs when f changes. Hence, the total volatility risk due to f changing is given by:

$$d\sigma/df = \partial\sigma/\partial f + (\rho v/f^\beta) \partial\sigma/\partial\alpha$$

This term also makes the delta hedge relatively insensitive to the particular value of β chosen, especially

near the ATM point.



Why is the sensitivity to beta reducing with the updated delta? Adjusted delta hedges are not overly sensitive to which fraction of the skew/smile arises from deterministic changes in the local volatility, and which arise from average changes in the local volatility in the adjusted delta (a division dictated to a large extent by beta)

Vega

1- Vega in Original SABR

Vega risk is the change in value when α changes by a unit amount

$$\frac{\partial V_{call}}{\partial \alpha} = \frac{\partial BS}{\partial \sigma_B} \cdot \frac{\partial \sigma_B(K, f; \alpha, \beta, \rho, v)}{\partial \alpha}.$$

Vega, in practice, can be found by bumping the instantaneous variance to $\alpha + \delta$, and then subtracting to determine the risk. It is traditional to scale vega so that it represents the change in value when the ATM volatility changes by a unit amount. Since $\delta \sigma_{ATM} = (\partial \sigma_{ATM} / \partial \alpha) \delta \alpha$, the vega risk is

$$\text{vega} \equiv \frac{\partial V_{call}}{\partial \alpha} = \frac{\partial BS}{\partial \sigma_B} \cdot \frac{\frac{\partial \sigma_B(K, f; \alpha, \beta, \rho, v)}{\partial \alpha}}{\frac{\partial \sigma_{ATM}(f; \alpha, \beta, \rho, v)}{\partial \alpha}}$$

To leading order, $\partial \sigma_B / \partial \alpha \approx \sigma_B / \alpha$ and $\partial \sigma_{ATM} / \partial \alpha \approx \sigma_{ATM} / \alpha$, so the vega risk is roughly given by

$$\text{vega} \approx \frac{\partial BS}{\partial \sigma_B} \cdot \frac{\sigma_B(K, f)}{\sigma_{ATM}(f)} = \frac{\partial BS}{\partial \sigma_B} \cdot \frac{\sigma_B(K, f)}{\sigma_B(f, f)}.$$

the volatility curve to calculate the total vega risk of a book of options.

As in the case of delta, the original SABR vega above ignores that, in theory, due to the correlation between f and alpha, changes in the instantaneous vol should result in changes in f . We share the Bartlett adjusted vega bellow to account for this effect.

2- Updated Bartlett Vega to account for changes in instantaneous vol caused by α

As in the case of delta, the vega risk should be calculated from the scenario:

$$f \rightarrow f + \delta_\alpha f, \alpha \rightarrow \alpha + \Delta\alpha$$

where $\delta_\alpha f$ is the average change in f caused by the change in alpha. It can be shown that:

$$\delta_\alpha f = (\rho f^\beta / \nu) * \Delta\alpha$$

Hence, the new vega risk is given by

$$\Lambda = \frac{\partial B}{\partial \sigma} \left(\frac{\partial \sigma}{\partial \alpha} + \frac{\partial \sigma}{\partial f} \frac{\rho f^\beta}{\nu} \right)$$

where the first term is the classic SABR vega, and the second term captures how changes in f (induced by change in alpha) impacts the local vol. As in the case of the Bartlett Delta, an option that has the Classic SABR delta + vega hedged also has the Bartlett vega hedged (as it is hedged against changes in local vol).

Vanna and Volga

$$\text{vanna} = \frac{\partial V_{call}}{\partial \rho} = \frac{\partial BS}{\partial \sigma_B} \cdot \frac{\partial \sigma_B(K, f; \alpha, \beta, \rho, \nu)}{\partial \rho},$$

$$\text{volga} = \frac{\partial V_{call}}{\partial \nu} = \frac{\partial BS}{\partial \sigma_B} \cdot \frac{\partial \sigma_B(K, f; \alpha, \beta, \rho, \nu)}{\partial \nu}.$$

Since ρ and ν are determined by fitting the implied volatility curve observed in the marketplace, the SABR model has risks to ρ and ν changing. Vanna basically expresses the risk to the skew increasing, and volga expresses the risk to the smile becoming more pronounced. These risks are easily calculated by using finite differences on the formula for σ_B in the SABR equations, and can be hedged by buying and selling OTM options.

Categories: [MARKETS](#) [OLD ARTICLES](#)



1 Comment



Wdowiak · 20 December 2022 at 0:17

Thank you so much

 REPLY

Leave a Reply

Name *

Email *

Website

What's on your mind?

Related Posts



MARKETS

Tariff Man – Season II

Download PDF Disclaimer: This article is intended solely for informational purposes. Bocconi Students Investment Club does not endorse any political candidate or party. Introduction Millions of Americans this week have cast their ballot for the [Read more...](#)

**MARKETS****Backtesting Series – Episode 2: Cross-Validation techniques**

Download PDF Introduction As announced in our previous episode, the second article in our Backtesting series will discuss Cross Validation and introduce a new Backtesting strategy, first presented by Giuseppe Paleologo (2024). If you haven't [Read more...](#)



MARKETS

Sports Betting Part 2: Further Exploring Cross-Bookie Arbitrage using Kelly Criterion and Odds Models

Download PDF Introduction In our previous exploration of sports betting, we examined basic market efficiency and evaluated whether straightforward strategies could uncover statistical arbitrage opportunities. Our findings showed that, while betting markets may appear exploitable [Read more...](#)

[instagram-feed]

Follow Us



COOKIE POLICY

Hestia | Developed by Themelsle