On the joint dynamics of the spot and the implied volatility surface

Sofiene El Aoud [†*†]   Frédéric Abergel [‡§]

December 8, 2014

Abstract

In this paper, we revisit the "Smile Dynamics" problem. In a previous work, Bergomi built a class of linear stochastic volatility models in which he specified the joint dynamics between the underlying and its instantaneous forward variances. The author introduced a quantity, which he called the Skew Stickiness ratio, in order to relate two quantities of interest: the first quantity is the correlation between the increments of the at-the-money implied volatility of maturity T and the log-returns of the underlying, while the second quantity is the implied skew of the same maturity T. In our work, we continue the study of the Skew stickiness ratio both from theoretical and empirical point of view. First, we provide a method to estimate the SSR (skew stickiness ratio) from option prices, this measure is called the implied SSR as it is conducted under the risk-neutral pricing measure $\mathcal{Q}$. Next to that, we recall how to measure the realized SSR under the real-world probability measure $\mathcal{P}$ and we point out empirically that there is a discrepancy between the implied SSR and the realized SSR. The empirical study shows also that the implied SSR, in the limit of short maturities, can take a value superior to 2 which is in discordance with the results obtained in linear stochastic volatility models. For this reason, we show that the positive quantity $(SSR_{Implied} - 2)$ is coherent with the presence of jumps in a stochastic volatility model.

[*]sofiene.elaoud@ecp.fr

[†]Exane Derivatives, 16 avenue Matignon, 75008 Paris

[‡]Ecole Centrale Paris, Laboratoire de Mathématiques Appliquées aux Systèmes, Grande Voie des Vignes, 92290 Châtenay Malabry, France

[§]frederic.abergel@ecp.fr

JEL Classification: JEL: C51 Model construction and estimation, JEL: C52 Model evaluation and testing.

Keywords: Leverage effect, realized skew, SSR indicator, option trading, arbitrage opportunity.

1 Introduction

The leverage effect is a well known feature in the equity markets, it has been deeply studied and documented by several authors (see [Bekaert and Wu(2000)], [Bouchaud et al.(2001)Bouchaud, Matacz, and Potters], [Bollerslev et al.(2006)Bollerslev, Litvinova, and Tauchen], [Ciliberti et al.(2009)Ciliberti, Bouchaud, and Potters]). This feature consists in the increase of the volatility following negative returns of the underlying, which explains the negative skewness of the underlying returns. The leverage effect influences considerably the derivatives prices. Indeed, the volatility smile in the equity option market has the particularity to be skewed, which means that the implied volatility is a decreasing function of the strike. Furthermore, the increment of the at-the-money volatility is negatively correlated with the log-return of the underlying, and this property is called "the implied leverage effect". The understanding of this property is a subject of interest since it is crucial to perform an efficient option delta-hedging strategy.

In [Bergomi(2009)], the author showed that the rate of decay of the at-the-money forward skew is linked to the covariance between the at-the-money volatility increment and the underlying log-return. Thus, he introduced a quantity called the Skew Stickiness Ratio (SSR) in order to quantify the relation between these two features. In this paper, we build on the work done in [Bergomi(2009)] and we conduct a study in order to understand the information contained in the Skew Stickiness Ratio. In the second section, we propose a model-free approach in order to estimate the SSR implied by option prices under the risk-neutral probability measure $\mathcal{Q}$. Following that, we carry out in the third section an empirical study which aims to compare the realized SSR under the historic probability measure $\mathcal{P}$ with the implied SSR under $\mathcal{Q}$. We define also an arbitrage strategy which enables to monetize the discrepancy between these two quantities. In the fourth section, we recall that, in the framework of linear stochastic volatility models, the implied SSR tends to the value 2 in the limit of short maturities. Since this property is contradictory with empirical findings, we introduce a stochastic volatility model with jumps in order to justify that the implied SSR can exceed the value 2 in the case of short maturities.

2 A model-free approach for implied Skew Stickiness Ratio estimation

Let $\sigma_{BS,t}(K,T)$ denote the implied volatility at time t for strike K and residual maturity T. We focus in this study on short maturity options ($0 < T \ll 1$). Thus, we can assume that the risk-free interest rate is null ($r = 0$). We suppose that the implied volatility, can be approximated for near the money strikes, by a quadratic function of the log-moneyness

$\log(\frac{K}{S_t})$:

$$\sigma_{BS,t}(K,T) \;=\; \sigma_{ATM,t}(T) + \mathcal{S}_T \log(\frac{K}{S_t}) + \mathcal{C}_T \log(\frac{K}{S_t})^2 + o\left(\log(\frac{K}{S_t})^2\right). \tag{2.1}$$

The Skew-Stickiness ratio, introduced by Bergomi in [Bergomi(2009)], is a function of the maturity T and is defined as:

$$R_T = \frac{d\left\langle \sigma_{ATM}(T), \log(S)\right\rangle_t}{\mathcal{S}_T d\left\langle \log(S)\right\rangle_t}. \tag{2.2}$$

The quantity R_T can be interpreted using the linear regression of the daily increments of the ATM volatility with maturity T on the daily log returns of the underlying:

$$d\sigma_{ATM,t}(T) \;=\; R_T \mathcal{S}_T d\log(S_t) + \eta dZ_t, \tag{2.3}$$

where Z is a Brownian motion such that $d\left\langle Z, \log(S)\right\rangle_t = 0$. This implies:

$$d\sigma_{BS,t}(K,T) \;=\; \mathcal{C}_T d\left\langle \log(S)\right\rangle_t + \left((R_T-1)\,\mathcal{S}_T + 2\mathcal{C}_T \log(\frac{S_t}{K})\right) d\log(S_t) + \eta dZ_t + O\left(\log(\frac{K}{S_t})^2\right).$$

and then:

$$d\sigma_{BS,t}(K,T) \;=\; \mathcal{C}_T d\left\langle \log(S)\right\rangle_t + (R_T-1)\mathcal{S}_T d\log(S_t) + \eta dZ_t + O\left(\log(\frac{K}{S_t})\right).$$

If $R_T = 1$ then $E(d\sigma_{BS,t}(K,T)|d\log(S_t)) = O(\mathcal{C}_T dt)$, which means that the sticky strike rule holds in average when $\mathcal{C}_T \sim 0$. On the other hand, if $R_T = 0$ then $E(d\sigma_{BS,t}(K,T)|d\log(S_t)) = -\mathcal{S}_T d\log(S_t) + O(\mathcal{C}_T dt)$. Thus, for $\mathcal{C}_T \sim 0$, the sticky-delta rule, where the implied volatility depends only on the moneyness, holds in average.

We aim here to estimate the value of R_T implied by option prices with maturity T without using any assumptions on the dynamics of the spot process. All we need here is to find the adequate value of R_T which solves the pricing equation of the option with maturity T.

Let $C(t, S_t, K, T, \sigma_{ATM,t}(T-t), \mathcal{S}_{T-t}, \mathcal{C}_{T-t}) = E^{\mathcal{Q}}(e^{-r(T-t)}(S_T - K)^+|\mathcal{F}_t)$ be the price at date t of an European call option with residual maturity $(T-t)$ and strike K. Applying Itô's Lemma yields:

$$dC_t = \frac{\partial C}{\partial t}dt + \frac{\partial C}{\partial S}dS_t + \frac{1}{2}\frac{\partial^2 C}{\partial S^2}d\left\langle S\right\rangle_t + \frac{\partial C}{\partial \sigma_{ATM}}d\sigma_{ATM,t}$$
$$+ \frac{1}{2}\frac{\partial^2 C}{\partial \sigma_{ATM}^2}d\left\langle \sigma_{ATM}\right\rangle_t + \frac{\partial^2 C}{\partial S \partial \sigma_{ATM}}d\left\langle S, \sigma_{ATM}\right\rangle_t.$$

The implied volatility corresponding to this option is equal to $\sigma_{BS}(K, T-t)$, and then:

$$C(t, S_t, K, T-t, \sigma_{ATM,T-t}, \mathcal{S}_{T-t}, \mathcal{C}_{T-t}) = P_{BS}(t, S_t, K, T, \sigma_{BS}(K, T-t)).$$

It is supposed that $r = 0$, which is an admissible hypothesis here since $(T-t)$ is small. It follows:

$$\frac{\partial P_{BS}}{\partial t} + \frac{1}{2}\sigma^2_{BS,t}(K, T-t) S_t^2 \frac{\partial^2 P_{BS}}{\partial S^2} \quad = \quad 0.$$

Let the hypothesis $\mathcal{H}^*$ be defined as follows:

- $\mathcal{H}^*$: For $T-t \sim 0$, the quantities $\sigma_{ATM}, \mathcal{S}, \mathcal{C}$ don't depend on $(T-t)$, and then $\sigma_{BS}(K, T-t)$ has no time dependence.

If the hypothesis $\mathcal{H}^*$ is satisfied then $\frac{\partial P_{BS}}{\partial t} = \frac{\partial C}{\partial t}$, and:

$$dC_t = -\frac{1}{2}\sigma^2_{BS,t}(K, T-t) S_t^2 \frac{\partial^2 P_{BS}}{\partial S^2} dt + \frac{\partial C}{\partial S} dS_t + \frac{1}{2}\frac{\partial^2 C}{\partial S^2} d\langle S\rangle_t + \frac{\partial C}{\partial \sigma_{BS}} d\sigma_{ATM,t}$$
$$+ \frac{1}{2}\frac{\partial^2 C}{\partial \sigma^2_{ATM}} d\langle \sigma_{ATM}\rangle_t + \frac{\partial^2 C}{\partial S \partial \sigma_{ATM}} d\langle S, \sigma_{ATM}\rangle_t .$$

In order to simplify the expressions, the following quantities x, τ and d_1 are introduced:

$$x = \log(\frac{K}{S_t}) \quad , \quad \tau = T - t \quad , \quad d_{1,t} = \frac{-x + \frac{1}{2}\sigma^2_{BS,t}(K,\tau)\tau}{\sigma_{BS,t}(K,\tau)\sqrt{\tau}}.$$

As explained in appendix (6.1), for near the money options ($x \sim 0$), high order derivatives of the option price C can be approximated using Black-Scholes greeks. Indeed, the second order partial derivative of C with respect to the spot price S writes:

$$\frac{\partial^2 C}{\partial S^2} \quad = \quad \frac{n(d_1)}{\sigma_{ATM} S_t \sqrt{\tau}} \left(1 - 3x\alpha + x^2(5\alpha^2 - \frac{5}{2}\beta)\right) + O\big(x^3\big) + O\big(\sqrt{\tau}\big) . \tag{2.4}$$

where α and β are functions of σ_{ATM} and defined as below:

$$\alpha(\sigma_{ATM}) = \frac{\mathcal{S}_{T-t}}{\sigma_{ATM}} \quad , \quad \beta(\sigma_{ATM}) = \frac{2\mathcal{C}_{T-t}}{\sigma_{ATM}}.$$

In addition, the second order partial derivative of C with respect to σ_{ATM} can be approximated using $d_{1,t}$, σ_{ATM}, x and τ:

$$\frac{\partial^2 C}{\partial \sigma^2_{ATM}} \quad = \quad \frac{S_t n(d_{1,t})}{\sigma^3_{ATM,t}\sqrt{\tau}} x^2 + O\big(x^3\big) + O\big(\sqrt{\tau}\big) , \tag{2.5}$$

and finally the approximation of $\frac{\partial^2 C}{\partial S \partial \sigma_{ATM}}$ can be given as:

$$\frac{\partial^2 C}{\partial S \partial \sigma_{ATM}} \quad = \quad \frac{n(d_{1,t})}{\sigma^2_{ATM,t}\sqrt{\tau}} \left(x - x^2(2\alpha - \sigma_{ATM,t}\alpha')\right) + O\big(\sqrt{\tau}\big) + O\big(x^3\big) . \tag{2.6}$$

Let K_L be a strike inferior to the spot value S_t, and K_H be the ATM strike ($K_H = S_t$). We define a portfolio X that contains, at time t, a quantity equal to (-1) unity of the option

C^L with strike K_L and a quantity equal to $(+n_{t,H})$ of the option C^H with strike K_H. The portfolio is delta-hedged continuously, and evolves as following:

$$dX_t = n_{t,H}\left(dC_t^H - \Delta_{t,H} dS_t\right) - \left(dC_t^L - \Delta_{t,L} dS_t\right)$$

Using (2.5) and (2.6), it can be deduced that the terms $\frac{\partial^2 C^H}{\partial \sigma_{ATM}^2}$ and $\frac{\partial^2 C^H}{\partial S \partial \sigma_{ATM}}$ are at order 1 in $\sqrt{\tau}$, and then can be neglected for $\tau \sim 0$.
In order to cancel the spot gamma sensitivity of the portfolio, the quantity $n_{t,H}$ is chosen as follows:

$$n_{t,H} = \frac{\frac{\partial^2 C^L}{\partial S^2}}{\frac{\partial^2 C^H}{\partial S^2}}.$$

Thus, the portfolio X has the following dynamics:

$$dX_t = \frac{1}{2} S_t^2 \left(\sigma_{BS}^2(K_L, T) \frac{\partial^2 P_{BS}^L}{\partial S^2} - n_{t,H} \sigma_{BS}^2(K_H, T) \frac{\partial^2 P_{BS}^H}{\partial S^2} \right) dt + \left(n_{t,H} \frac{\partial C^H}{\partial \sigma_{ATM}} - \frac{\partial C^L}{\partial \sigma_{ATM}} \right) d\sigma_{ATM,t}$$
$$- \frac{1}{2} \frac{\partial^2 C^L}{\partial \sigma_{ATM}^2} d \left\langle \sigma_{ATM} \right\rangle_t - \frac{\partial^2 C^L}{\partial S \partial \sigma_{ATM}} R_T \mathcal{S}_t S_t d \left\langle \log(S) \right\rangle_t .$$

In order to simplify the notations, we introduce the following quantities:

$$\Gamma_t = \frac{1}{2} S_t^2 \left(\sigma_{BS}^2(K_L, T) \frac{\partial^2 P_{BS}^L}{\partial S^2} - n_{t,H} \sigma_{BS}^2(K_H, T) \frac{\partial^2 P_{BS}^H}{\partial S^2} \right),$$
$$\vartheta_t = n_{t,H} \frac{\partial C^H}{\partial \sigma_{ATM}} - \frac{\partial C^L}{\partial \sigma_{ATM}}.$$

Under the risk-neutral probability measure $\mathcal{Q}$, we have $E(dX_t) = rX_t dt = 0$, then:

$$\frac{\partial^2 C^L}{\partial S \partial \sigma_{ATM}} R_T^{\mathcal{Q}} S_t \mathcal{S}_T d \left\langle \log(S) \right\rangle_t = \Gamma_t dt + \vartheta_t E^{\mathcal{Q}}(d\sigma_{ATM,t}) - \frac{1}{2} \frac{\partial^2 C^L}{\partial \sigma_{ATM}^2} d \left\langle \sigma_{ATM} \right\rangle_t . \quad (2.7)$$

Based on (2.3), it can be deduced that $d \left\langle \sigma_{ATM} \right\rangle_t = R_T^2 \mathcal{S}_T^2 d \left\langle \log(S) \right\rangle_t + \eta^2 d \left\langle Z \right\rangle_t$. Thus, we obtain the following equation:

$$\frac{1}{2} \frac{\partial^2 C^L}{\partial \sigma_{ATM}^2} (R_T^{\mathcal{Q}})^2 \mathcal{S}_T^2 d \left\langle \log(S) \right\rangle_t + R_T^{\mathcal{Q}} \left(\frac{\partial^2 C^L}{\partial S \partial \sigma_{ATM}} S_t \mathcal{S}_T d \left\langle \log(S) \right\rangle_t - \vartheta_t \mathcal{S}_T E^{\mathcal{Q}}(d \log(S_t)) \right)$$
$$- \left(\Gamma_t - \frac{1}{2} \frac{\partial^2 C^L}{\partial \sigma_{ATM}^2} \eta^2 \right) dt = 0.$$

Under the simplifying assumptions that $\eta \sim 0$ and $E^{\mathcal{Q}}(d \log(S_t)) = 0$, the last equation becomes:

$$\frac{1}{2} \frac{\partial^2 C^L}{\partial \sigma_{ATM}^2} (R_T^{\mathcal{Q}})^2 \mathcal{S}_T^2 d \left\langle \log(S) \right\rangle_t + \frac{\partial^2 C^L}{\partial S \partial \sigma_{ATM}} S_t \mathcal{S}_T R_T^{\mathcal{Q}} d \left\langle \log(S) \right\rangle_t - \Gamma_t dt = 0. \quad (2.8)$$

The resolution of the equation (2.8) enables to obtain the value of $R_T^{\mathcal{Q}}$:

$$R_T^{\mathcal{Q}} = \frac{-\frac{\partial^2 C^L}{\partial S \partial \sigma_{ATM}} S_t \mathcal{S}_T d\left\langle \log(S) \right\rangle_t + \sqrt{D_t}}{\frac{\partial^2 C^L}{\partial \sigma_{ATM}^2} (\mathcal{S}_T)^2 d\left\langle \log(S) \right\rangle_t}, \tag{2.9}$$

where D_t represents the discriminant of the quadratic equation (2.8):

$$D_t = \left(\frac{n(d_1^L) x_L}{\sigma_{ATM}^2 \sqrt{\tau}} \right)^2 S_t^2 \mathcal{S}_T^2 d\left\langle \log(S) \right\rangle_t^2 + 2\Gamma_t dt \left(\frac{S_t n(d_1^L)}{\sigma_{ATM}^3 \sqrt{\tau}} x_L^2 \right) \mathcal{S}_T^2 d\left\langle \log(S) \right\rangle_t,$$

3 Arbitraging the Skew Stickiness Ratio

Let $R_T^{\mathcal{P}}$ be the value of the quantity R_T under the historic probability measure $\mathcal{P}$, $R_T^{\mathcal{P}}$ can be determined using the linear regression of the daily increments of the ATM volatility with maturity T on the daily log-returns of the spot process S:

$$d\sigma_{ATM,t}(T) = R_T \mathcal{S}_T d\log(S_t) + \eta dZ_t,$$

then $R_T^{\mathcal{P}}$ can be estimated as follows:

$$R_{t,T}^{\mathcal{P}} = \frac{\sum_{i=t-L}^{t} \left(\sigma_{ATM,i}(T) - \sigma_{ATM,i-1}(T)\right)\left(\log(S_i) - \log(S_{i-1})\right)}{\mathcal{S}_{t,T}\left(\sum_{i=t-L}^{t} \left(\log(S_i) - \log(S_{i-1})\right)^2\right)}.$$

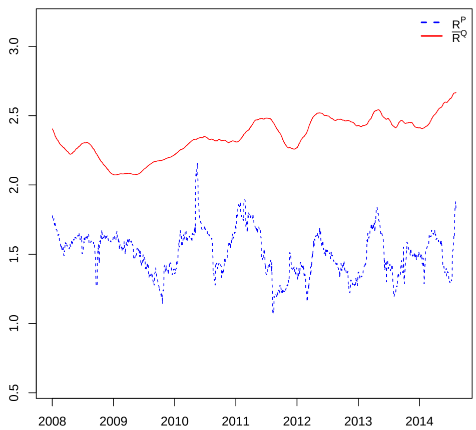
In order to to make this measure less noisy, $R_{t,T}^{\mathcal{P}}$ is computed in the the empirical study as follows:

$$R_{t,T}^{\mathcal{P}} = \frac{\sum_{i=t-L}^{t} \left(\sigma_{ATM,i}(T) - \sigma_{ATM,i-1}(T)\right)\left(\log(S_i) - \log(S_{i-1})\right)}{\left(\frac{1}{L+1}\sum_{i=t-L}^{t} \mathcal{S}_{i,T}\right)\left(\sum_{i=t-L}^{t} \left(\log(S_i) - \log(S_{i-1})\right)^2\right)}, \tag{3.1}$$

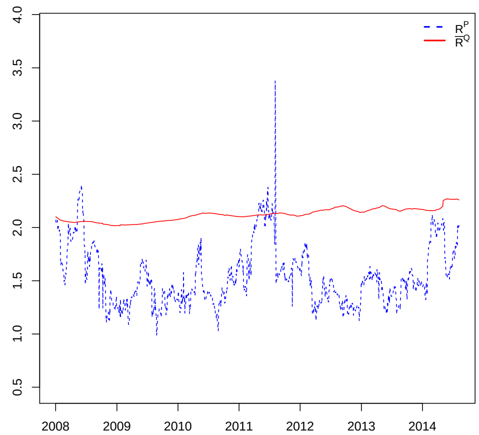
We study the historic evolution of $R_T^{\mathcal{P}}$ and $R_T^{\mathcal{Q}}$ and investigate the presence of discrepancies between these two quantities.

3.1 Empirical study

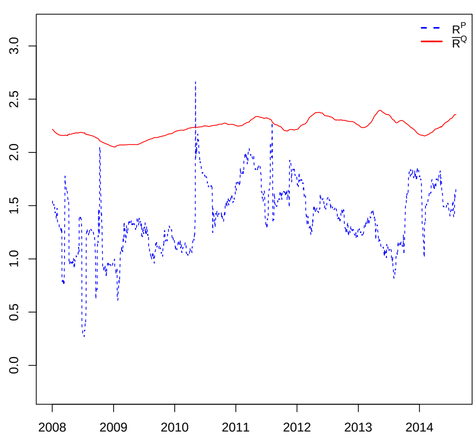
We carry out an empirical study on the skew stickiness ratio using the results provided in (2.9) and (3.1). We use historical data from option markets and from underlying price series, and we conduct the study on several assets including the SPX index, SX5E index, Financial Select Sector (XLF), Technology Select Sector (XLK). We focus here on $3M$ options ($T = \frac{3}{12}$) which is consistent with the hypothesis of short maturity ($T \sim 0$). The hypothesis $\mathcal{H}^*$ is supposed to be satisfied, and the quantity $R_T^{\mathcal{Q}}$ is obtained through the resolution of the equation (2.8). We dispose of data ranging from 01/01/2007 to 07/08/2014. For every day i, we estimate $R_{i,T}^{\mathcal{Q}}$ and $R_{i,T}^{\mathcal{P}}$ solutions of (2.9) and (3.1) respectively. In order to have two quantities calculated on the same window of data and then easily comparable, we introduce the quantity $\bar{R}_{i,T}^{\mathcal{Q}} = \sum_{j=i-L}^{i} R_{j,T}^{\mathcal{Q}}$. The following graphs give the evolution of $\bar{R}_{i,T}^{\mathcal{Q}}$ and $R_{i,T}^{\mathcal{P}}$ using the window parameter $L = 50$ Days:



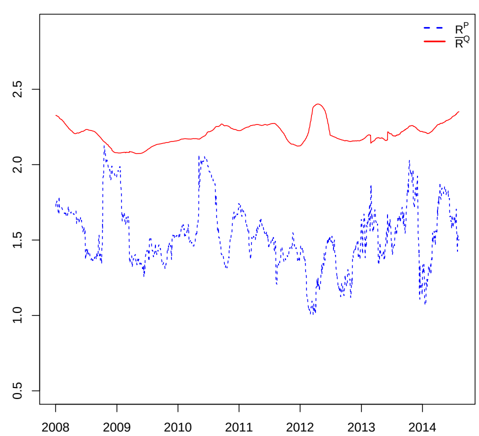
(a) SPX



(b) XLF



(c) XLK



(d) SX5E

The inspection of the graphs above show that the quantity $\bar{R}_T^Q$ can be significantly larger than the value of 2. This finding is contradictory with the characteristics of linear stochastic volatility models. Indeed, Bergomi demonstrated in [Bergomi(2009)], that in the setting of linear models, the quantity $\bar{R}_T^Q$ converges to the value 2 in the limit of small maturities. In addition, he showed that, in the case of time-homogeneous models and a flat term-structure of variance, $\bar{R}_T^Q$ is restricted to the interval [1, 2]. This discrepancy was pointed out by the authors in [Vargas et al.(2013)Vargas, Dao, and Bouchaud] who justified it by the existence of a non-linear leverage effect in equity markets which can not be captured using linear stochastic volatility models. Thus, the empirical study confirms the importance of use of non-linear models on some assets like the SPX index for which the average value of $\bar{R}_T^{\mathcal{Q}}$ is the highest among the other examples.

In addition to that, it can be noticed that $R_T^{\mathcal{P}}$ is more oscillatory and unstable compared to $\bar{R}_T^{\mathcal{Q}}$, and that there is generally a discrepancy between the values of these two quantities. Bergomi pointed out in [Bergomi(2009)] that $R_T^{\mathcal{P}}$ is inferior to 2 which is, in the setting of linear models, the limit of $R_T^{\mathcal{Q}}$ when T tends to 0. This is true in average (but not always) and raises the question of arbitraging the spread $(R_T^{\mathcal{Q}} - R_T^{\mathcal{P}})$. In [Bergomi(2009)], the author tried to establish a trading strategy whose $P\&L$ is proportional to the spread $(2 - R_T^{\mathcal{P}})$. In the next paragraph, we build on this work and we define a different strategy which aims to take profit from the discrepancy between $R_T^{\mathcal{Q}}$ and $R_T^{\mathcal{P}}$.

3.2 Taking advantage of the Skew Stickiness Ratio discrepancy

Under the real-world probability measure $\mathcal{P}$, the portfolio X, defined previously, has the following dynamics:

$$\begin{aligned} dX_t = \Gamma_t dt + &\left(n_{t,H}\frac{\partial C^H}{\partial \sigma_{ATM}} - \frac{\partial C^L}{\partial \sigma_{ATM}}\right) d\sigma_{ATM,t} \\ &- \frac{1}{2}\frac{\partial^2 C^L}{\partial \sigma_{ATM}^2}\left((R_T^{\mathcal{P}})^2 \mathcal{S}_T^2 d\left\langle \log(S)\right\rangle_t + \eta^2 dt\right) - \frac{\partial^2 C^L}{\partial S \partial \sigma_{ATM}} R_T^{\mathcal{P}} \mathcal{S}_T S_t d\left\langle \log(S)\right\rangle_t . \end{aligned}$$

Using the definition of $R_T^{\mathcal{Q}}$ given in (2.7), it can be deduced that:

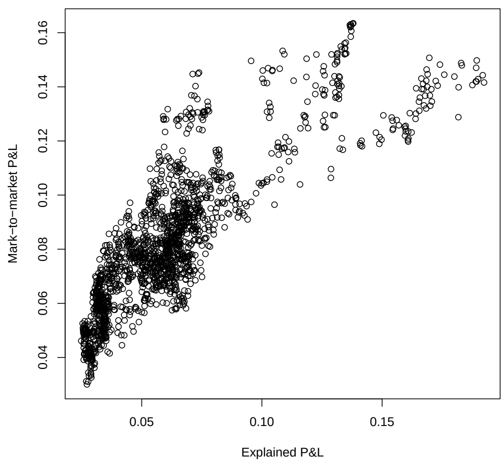
$$\begin{aligned} dX_t = &\left(n_{t,H}\frac{\partial C^H}{\partial \sigma_{ATM}} - \frac{\partial C^L}{\partial \sigma_{ATM}}\right)\left(d\sigma_{ATM,t} - E^{\mathcal{Q}}\left(d\sigma_{ATM,t}\right)\right) \\ &+ \frac{1}{2}\frac{\partial^2 C^L}{\partial \sigma_{ATM}^2}\left((R_T^{\mathcal{Q}})^2 - (R_T^{\mathcal{P}})^2\right)\mathcal{S}_T^2 d\left\langle \log(S)\right\rangle_t + \frac{\partial^2 C^L}{\partial S \partial \sigma_{ATM}}\left(R_T^{\mathcal{Q}} - R_T^{\mathcal{P}}\right)\mathcal{S}_T S_t d\left\langle \log(S)\right\rangle_t . \end{aligned}$$

Then:

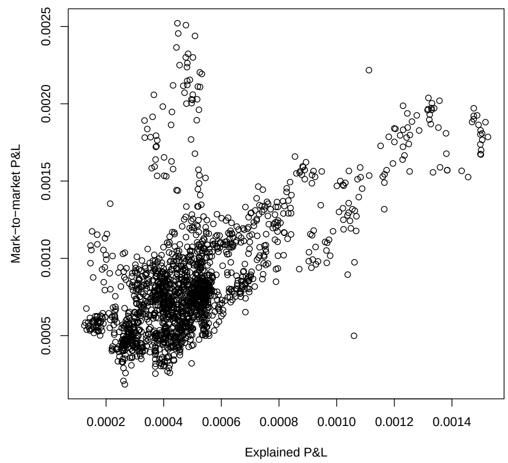
$$E(dX_t) \quad = \quad \left(R_T^{\mathcal{Q}} - R_T^{\mathcal{P}}\right)\left(\frac{1}{2}\frac{\partial^2 C^L}{\partial \sigma_{ATM}^2}\left(R_T^{\mathcal{Q}} + R_T^{\mathcal{P}}\right)\mathcal{S}_T^2 + \frac{\partial^2 C^L}{\partial S \partial \sigma_{ATM}}\mathcal{S}_T S_t\right) d\left\langle \log(S)\right\rangle_t$$

which is a function of the spread $\left(R_T^{\mathcal{Q}} - R_T^{\mathcal{P}}\right)$.

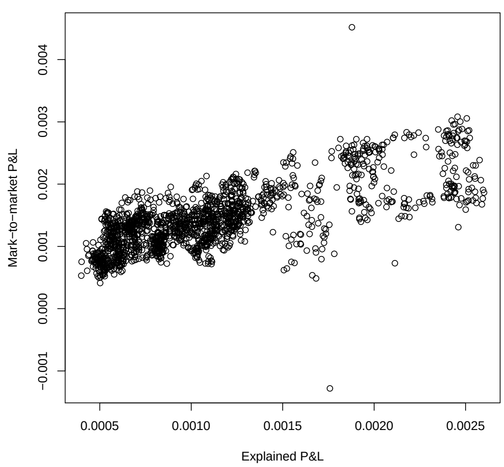
In the following, we run a backtest of a strategy which replicates the portfolio X. The strategy consists in selling every day a 3 months option with moneyness 95%, buying the quantity n_H of options with moneyness 100%, and doing the necessary delta-hedging. The portfolio is unwound the next day and started again. The scatter plots show the daily mark-to-market $P\&L$ of the portfolio as a function of the theoretical $P\&L$ given by the analytic expression of dX_t.



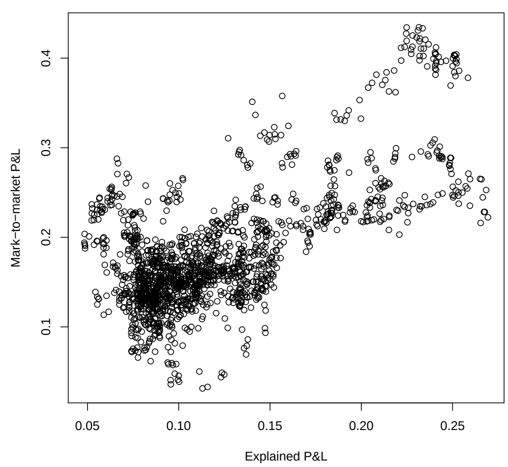
(e) SPX



(f) XLF



(g) XLK



(h) SX5E

Statistics on PnL explanation	Regression slope	R-squared
SPX	0.704	0.699
SX5E	0.932	0.462
XLF	1.02	0.495
XLK	0.694	0.546

The scatter plots above and the R^2 of the regressions show how well the mark-to-market $P\&L$ matches the theoretical $P\&L$. It should be pointed out here that the difference between these two quantities can be explained by several factors. Indeed, the quantities $\frac{\partial^2 C}{\partial S \partial \sigma_{ATM}}$, $\frac{\partial^2 C}{\partial S^2}$ and $\frac{\partial^2 C}{\partial \sigma_{ATM}^2}$ are not exact but approximated at order 2 in x and 0 in τ, besides the hypothesis

$\mathcal{H}^*$ is assumed to be verified, which may be a source of noise in the $P\&L$ when it's not the case.

4 Limit of the Skew Stickiness Ratio for short maturities

It was clear through the empirical study that R_T^Q can exceed the value of 2 when $T \sim 0$. This empirical finding is in discordance with the theoretical results established by Bergomi in [Bergomi(2009)] for the class of linear stochastic volatility models. In order to explain this phenomenon, the authors in [Vargas et al.(2013)Vargas, Dao, and Bouchaud] proposed an asymmetric Garch model which accounts for the non-linear leverage effect in the equity market. This model can produce values of R_T^Q superior to 2 for a maturity T of the order of several days. Meanwhile, for a 3 Months maturity ($T = 0.25$), which is the maturity considered in the empirical study, the quantity R_T^Q produced by the asymmetric Garch model can not be superior to 2. This is due to the fact that the implied at-the-money skew nearly coincides with the quantity $\frac{Skewness_T}{6\sqrt{T}}$ when T is of the order of several months (it can be precised here that $Skewness_T$ represents the skewness of $\log(\frac{S_T}{S_0})$). Consequently, the asymmetric Garch model doesn't justify theoretically the empirical observation $R_T^Q > 2$ for $T = 0.25$.

In order to support theoretically this empirical observation, we propose here a model which enables to obtain a value of R_T^Q superior to 2 when $T \sim 0$.

We suppose that, under the risk-neutral probability measure Q, the spot process S has the following dynamics:

$$\frac{dS_t}{S_t} = (r - \lambda k)dt + \sigma_t\sqrt{1-\rho_t^2}dW_t^{(1)} + \sigma_t\rho_t dW_t^{(2)} + (J_t - 1)dN_t$$

where N_t is a Poisson process with intensity λ, r is the risk-free rate, J_t is a random positive variable and $k = E(J_t - 1)$. The instantaneous volatility $\sigma_t = \sigma(Y_t)$ is a deterministic function of the stochastic process Y which evolves as follows:

$$dY_t = b_t dW_t^{(2)} \qquad (4.1)$$

where $d\left\langle W^{(1)}, W^{(2)}\right\rangle_t = 0$.
It can be mentioned here that b_t and ρ_t can also depend on Y_t, that is:

$$b_t = b(Y_t) \quad , \quad \rho_t = \rho(Y_t)$$

and:

$$\int_0^t b_s^2 ds < \infty$$

Let $P_t = e^{-r(T-t)}E^Q((S_T - K)^+)$ be the price of an European call option on S with strike K and maturity T. Under the assumption that b and λ are small, the following expansion for the option price P can be made:

$$P = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} b^i \lambda^j P_{i,j}$$

Through the use of a singular perturbation technique in b and λ, the option price P can be approximated at order 1 in b and λ by $\hat{P}$ as proved in (6.2):

$$\hat{P}(t, S_t, y_t) = P_{0,0} + bP_{1,0} + \lambda P_{0,1}, \tag{4.2}$$

where:

$$P_{0,0}(t, S_t, y_t) = P_{BS}(t, S_t, \sigma(y_t), K, T), \tag{4.3}$$

$$P_{1,0}(t, S_t, y_t) = \frac{T-t}{2} S_t \sigma(y_t)\sigma'(y_t)\rho_t \frac{\partial^2 P_{0,0}}{\partial S \partial \sigma}, \tag{4.4}$$

$$P_{0,1}(t, S_t, y_t) = (T-t)(\phi(P_{0,0}) - S_t k \frac{\partial P_{0,0}}{\partial S}), \tag{4.5}$$

where the operator ϕ is defined as:

$$\begin{aligned} \phi(P) &= E(P(t, S.J, y)) - P(t, S, y) \\ &= \int_0^{\infty} P(t, S \times j, y) f_J(j) dj - P(t, S, y), \end{aligned}$$

and f_J is the probability density function of the random variable J.

Let $I_t(K,T)$ be the implied volatility of P, that is $I_t(K,T)$ is defined such that $P_t^{K,T} = P_{BS}(t, S_t, K, T, I_t(K,T))$. In order to approximate $I_t(K,T)$ at first order in b and λ, a Taylor expansion of the implied volatility $I_t(K,T)$ is carried out in (6.3) and we prove that:

$$I_t(K,T) = \hat{I}_t(K,T) + O\big(b^2 + \lambda^2 + b\lambda\big),$$

where $\hat{I}_t(K,T)$ is defined as follows:

$$\hat{I}_t(K,T) = d(\rho, b, \lambda, T-t) + a(\rho, b, \lambda, T-t)\log(\frac{K}{F_{t,T}}) + c(\rho, b, \lambda, T-t)\log(\frac{K}{F_{t,T}})^2,$$

the quantities a, c and d are defined as follows:

$$\begin{aligned} a(\rho, b, \lambda, T-t) &= \frac{b_t \rho_t}{2}\frac{\sigma'(y)}{\sigma(y)} + \lambda_t \frac{E((J-1)^3)}{6\sigma^3(y)(T-t)} - \frac{\lambda_t E((J-1)^4)}{6\sigma^3(y)(T-t)}, \\ c(\rho, b, \lambda, T-t) &= \frac{\lambda_t E((J-1)^4)}{24\sigma^5(y)(T-t)^2}, \end{aligned}$$

and:

$$\begin{aligned} d(\rho, b, \lambda, T-t) = \sigma(y) + b_t\rho_t\sigma'(y)\frac{\sigma(y)}{4}(T-t) + \lambda_t \frac{E((J-1)^2)}{2\sigma(y)} - \lambda_t \frac{E((J-1)^3)}{4\sigma(y)} \\ + \frac{\lambda_t}{24} E((J-1)^4)(\frac{15}{4\sigma(y)} - \frac{1}{\sigma^3(y)(T-t)}) \end{aligned}$$

In the absence of jumps ($\lambda = 0$), the quantity $R_T^{\mathcal{Q}}$ writes:

$$R_T^{\mathcal{Q}} = \frac{\sigma'(y_t)\sigma(y_t)b_t\rho_t + \frac{b_t\rho_t}{4}(T-t)\left(\sigma'(y_t)^2 + \sigma(y_t)\sigma''(y_t)\right)\sigma(y_t)b_t\rho_t}{\frac{b_t\rho_t}{2}\sigma'(y_t)\sigma(y_t)},$$

then: $\underset{T\to 0}{lim} R_T = 2$.

Suppose now that the intensity of jumps is strictly positive ($\lambda > 0$). For the ATM skew $a(\rho, b, \lambda, T-t)$ not to be infinite when $T = t$, it can be supposed that $E((J-1)^3) = E((J-1)^4)$. Thus, $R_T^{\mathcal{Q}}$ has the following limit when $T \to 0$:

$$\underset{T\to 0}{lim} R_T = 2 + \frac{\lambda}{\sigma^2(y_t)}\left(\frac{E((J-1)^3)}{2} - E((J-1)^2)\right),$$

Consequently, according to this model, R_T can take a value superior to 2 when $T \to 0$.

5 Conclusion

In this chapter, we conducted a study on the Skew-Stickiness Ratio. We provided a model-free approach which allows to measure the SSR under the risk-neutral pricing measure $\mathcal{Q}$ for a given maturity T, this approach is free from any assumption on the dynamics of the underlying asset. We also showed that the historical value of the SSR under the real-world probability measure $\mathcal{P}$ can be different from its value under the pricing measure $\mathcal{Q}$ for short maturities. Thus, we suggested a trading strategy which aims to monetize the difference between $R_T^{\mathcal{Q}}$ and $R_T^{\mathcal{P}}$. We tested the suggested strategy on real data in order to show that the assumptions made in the theoretical study are not too strong and that Mark-to-market $P\&L$ stays well explained by the model. We focused then on the empirical observation that $R_T^{\mathcal{Q}}$ may exceed the value 2 in the limit of short maturities. Since the asymmetric Garch model fails to reproduce this result for $T = 0.25$, we proposed a stochastic volatility model with Jumps that can reproduce this result.

6 Appendix

6.1 Appendix 1

Suppose that:

$$\sigma_{BS}(K,T) = \sigma_{ATM}\left(1+\alpha(\sigma_{ATM})\log(\frac{K}{S_t})+\frac{1}{2}\beta(\sigma_{ATM})\left(\log(\frac{K}{S_t})\right)^2\right).$$

We have:

$$\frac{\partial^2 Q}{\partial S^2} = \frac{\partial^2 P_{BS}}{\partial S^2}+2\frac{\partial^2 P_{BS}}{\partial S\partial\sigma_{BS}}\frac{\partial\sigma_{BS}}{\partial S}+\frac{\partial^2 P_{BS}}{\partial\sigma_{BS}^2}(\frac{\partial\sigma_{BS}}{\partial S})^2+\frac{\partial P_{BS}}{\partial\sigma_{BS}}\frac{\partial^2\sigma_{BS}}{\partial S^2},$$

Then:

$$S_t^2\frac{\partial^2 Q}{\partial S^2} = \frac{S_t n(d_1)}{\sigma_{BS}\sqrt{\tau}}\left(1+2\sqrt{\tau}d_2\sigma_{ATM}(\alpha+\beta x)+d_1 d_2\tau\sigma_{ATM}^2(\alpha+\beta x)^2+\sigma_{BS}\tau\sigma_{ATM}(\alpha+x\beta+\beta)\right),$$

Recall that $d_1 = \frac{-x+\frac{1}{2}\sigma_{BS}^2\tau}{\sigma_{BS}\sqrt{\tau}}$ and $d_2 = d_1-\sigma_{BS}\sqrt{\tau}$, then the expression can be simplifies as follows:

$$\begin{aligned}
S_t^2\frac{\partial^2 Q}{\partial S^2} &= \frac{S_t n(d_1)}{\sigma_{BS}\sqrt{\tau}}\left(1-2x(\alpha+\beta x)(1-\alpha x-\frac{1}{2}\beta x^2)+x^2\alpha^2+o(x^3)\right)+O(\sqrt{\tau}),\\
&= \frac{S_t n(d_1)}{\sigma_{BS}\sqrt{\tau}}\left(1-2x\alpha(\sigma_{ATM})+x^2(3\alpha^2(\sigma_{ATM})-2\beta(\sigma_{ATM}))+o(x^3)\right)+O(\sqrt{\tau}),\\
&= \frac{S_t n(d_1)}{\sigma_{ATM}\sqrt{\tau}}\left(1-2x\alpha+x^2(3\alpha^2-2\beta)\right)\left(1-x\alpha-\frac{1}{2}\beta x^2\right)+o(x^3)+O(\sqrt{\tau}),\\
&= \frac{S_t n(d_1)}{\sigma_{ATM}\sqrt{\tau}}\left(1-3x\alpha(\sigma_{ATM})+x^2(5\alpha^2(\sigma_{ATM})-\frac{5}{2}\beta(\sigma_{ATM}))\right)+o(x^3)+O(\sqrt{\tau}),
\end{aligned}$$

The quantity $\frac{\partial^2 Q}{\partial\sigma_{ATM}^2}$ can also be written using partial derivatives of P_{BS}:

$$\frac{\partial^2 Q}{\partial\sigma_{ATM}^2} = \frac{\partial^2 P_{BS}}{\partial\sigma_{BS}^2}(\frac{\partial\sigma_{BS}}{\partial\sigma_{ATM}})^2+\frac{\partial P_{BS}}{\partial\sigma_{BS}}\frac{\partial^2\sigma_{BS}}{\partial\sigma_{ATM}^2}$$

Using the expressions of d_1 and d_2, we deduce that:

$$\begin{aligned}
\frac{1}{2}\sigma_{ATM}^2\frac{\partial^2 Q}{\partial\sigma_{ATM}^2} &= \frac{S_t n(d_1)d_1 d_2\sqrt{\tau}}{\sigma_{BS}}(\frac{\partial\sigma_{BS}}{\partial\sigma_{ATM}})^2+O(\sqrt{\tau}),\\
&= \frac{S_t n(d_1)\sigma_{ATM}^2}{2\sigma_{BS}^3\sqrt{\tau}}x^2+O(x^3)+O(\sqrt{\tau}),\\
&= \frac{S_t n(d_1)}{2\sigma_{ATM}\sqrt{\tau}}x^2+O(x^3)+O(\sqrt{\tau}),
\end{aligned}$$

Using the same method, we provide analytic approximation for $\frac{\partial^2 Q}{\partial S \partial \sigma_{ATM}}$ up to order 2 in x and 0 in τ. Indeed:

$$\frac{\partial^2 Q}{\partial S \partial \sigma_{ATM}} = \frac{\partial^2 P_{BS}}{\partial S \partial \sigma_{BS}} \frac{\partial \sigma_{BS}}{\partial \sigma_{ATM}} + \frac{\partial^2 P_{BS}}{\partial \sigma_{BS}^2} \frac{\partial \sigma_{BS}}{\partial \sigma_{ATM}} \frac{\partial \sigma_{BS}}{\partial S} + \frac{\partial P_{BS}}{\partial \sigma_{BS}} \frac{\partial^2 \sigma_{BS}}{\partial S \partial \sigma_{ATM}}.$$

Then:

$$\begin{aligned} S\sigma_{ATM} \frac{\partial^2 Q}{\partial S \partial \sigma_{ATM}} &= \frac{S_t n(d_1)}{\sigma_{ATM}\sqrt{\tau}} \left(\frac{\sigma_{ATM}^2}{\sigma_{BS}^2} x - \frac{\sigma_{ATM}^3}{\sigma_{BS}^3} x^2 (\alpha(\sigma_{ATM}) + x\beta(\sigma_{ATM})) \right) \frac{\partial \sigma_{BS}}{\partial \sigma_{ATM}} + O\big(\sqrt{\tau}\big), \\ &= \frac{S_t n(d_1)}{\sigma_{ATM}\sqrt{\tau}} \big((1 - 2\alpha(\sigma_{ATM})x)x - (1 - 3\alpha x)x^2(\alpha + x\beta) \big) \frac{\partial \sigma_{BS}}{\partial \sigma_{ATM}} + O\big(\sqrt{\tau}\big), \\ &= \frac{S_t n(d_1)}{\sigma_{ATM}\sqrt{\tau}} \big(x - x^2 (2\alpha(\sigma_{ATM}) - \sigma_{ATM}\alpha'(\sigma_{ATM})) \big) + O\big(\sqrt{\tau}\big) + O\big(x^3\big). \end{aligned}$$

6.2 Appendix 2

Let P be the price of an European option on the stock S with maturity T and payoff $H(S_T)$:

$$P(t, S_t, Y_t, K, T) = e^{-r(T-t)} E^{\mathcal{Q}}(H(S_T)|F_t).$$

Using It$\hat{o}$'s lemma:

$$\mathcal{L}P = 0,$$

where the operator $\mathcal{L}$ is defined as follows:

$$\mathcal{L}P = \frac{\partial P}{\partial t} + \frac{\partial P}{\partial S} S_t(r - \lambda k) + \frac{1}{2} S_t^2 \sigma^2(y) \frac{\partial^2 P}{\partial S^2} + \frac{1}{2} \frac{\partial^2 P}{\partial y^2} b^2 + \frac{\partial^2 P}{\partial S \partial y} S_t \sigma(y) b_t \rho_t + \lambda \phi(P) - rP.$$

where the operator ϕ is defined as:

$$\begin{aligned} \phi(P) &= E(P(t, S.J, y)) - P(t, S, y) \\ &= \int_0^\infty P(t, Sj, y) f_J(j) dj - P(t, S, y), \end{aligned}$$

and f_J is the probability density function of the random variable J. In order to separate the terms in $\mathcal{L}$ by their orders in b and λ, the following differential operators are introduced:

$$\begin{aligned} \mathcal{L}_{0,0} &= \frac{\partial}{\partial t} + r(S_t \frac{\partial}{\partial S_t} - .) + \frac{1}{2}\sigma^2(y) S_t^2 \frac{\partial^2}{\partial S_t^2}, \\ \mathcal{L}_{1,0} &= S_t \sigma(y) \rho_t \frac{\partial^2}{\partial S_t \partial y}, \\ \mathcal{L}_{2,0} &= \frac{1}{2} \frac{\partial^2}{\partial y^2}, \\ \mathcal{L}_{0,1} &= \phi(.) - k S_t \frac{\partial}{\partial S_t}, \end{aligned}$$

which implies:

$$\mathcal{L}P \;=\; \left(\mathcal{L}_{0,0} + b\mathcal{L}_{1,0} + b^2\mathcal{L}_{2,0} + \lambda\mathcal{L}_{0,1}\right) P = 0 \tag{6.1}$$

Under the assumption that b and λ are small, the option price P can be expanded in powers of b and λ:

$$P \;=\; \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} b^i \lambda^j P_{i,j} \tag{6.2}$$

By insertion of (6.2) in (6.1) and regroupment of terms by their powers in b and λ, the following equations are obtained:

$$\begin{aligned}(0,0) &: \mathcal{L}_{0,0}P_{0,0} = 0,\\ (1,0) &: \mathcal{L}_{0,0}P_{1,0} + \mathcal{L}_{1,0}P_{0,0} = 0,\\ (0,1) &: \mathcal{L}_{0,0}P_{0,1} + \mathcal{L}_{0,1}P_{0,0} = 0.\end{aligned}$$

Using the zero-order term $(0,0)$, it can be seen that $P_{0,0}$ is the solution of the equation:

$$\mathcal{L}_{BS}(\sigma(y))P_{0,0} \;=\; 0,$$

with the final condition:

$$P_{0,0}(T, S_T, Y_T) \;=\; H(S_T).$$

Then, $P_{0,0}$ is the Black-Scholes price of the option with implied volatility equal to $\sigma(y)$:

$$P_{0,0}(t, S_t, Y_t) = P_{BS}(t, S_t, \sigma(y_t), K, T).$$

The equation of the $(1,0)$-term (at order 1 in b and 0 in λ) shows that $P_{1,0}$ verifies:

$$\mathcal{L}_{BS}\left(\sigma(y)\right) P_{1,0} \;=\; -S_t\sigma(y)\rho_t \frac{\partial^2 P_{0,0}}{\partial S \partial y},$$

with the final condition:

$$P_{1,0}(T, S_T, Y_T) \;=\; 0.$$

In order to determine $P_{1,0}$, we start by computing the term $S_t\sigma(y)\frac{\partial^2}{\partial S\partial y}\mathcal{L}_{BS}(\sigma(y))$:

$$S_t\sigma(y)\frac{\partial^2}{\partial S\partial y}\mathcal{L}_{BS}(\sigma(y)) \;=\; \mathcal{L}_{BS}(\sigma(y))S_t\sigma(y)\frac{\partial^2}{\partial S\partial y} + \sigma^2(y)\sigma'(y)S_t\frac{\partial}{\partial S_t}(S_t^2\frac{\partial^2}{\partial S_t^2})$$

Let $R(t, S_t, y_t)$ be defined as follows:

$$R(t, S_t, y_t) \;=\; (T-t)S_t\sigma(y_t)\rho_t\frac{\partial^2 P_{0,0}}{\partial S\partial y} - \frac{(T-t)^2}{2}\sigma^2(y_t)\sigma'(y_t)\rho_t\mathcal{D}_{1,S}\mathcal{D}_{2,S}P_{0,0}$$

where $\mathcal{D}_{1,S}$ and $\mathcal{D}_{2,S}$ denote the two following differential operators:

$$\mathcal{D}_{1,S} = S_t \frac{\partial}{\partial S_t} \quad , \quad \mathcal{D}_{2,S} = S_t^2 \frac{\partial^2}{\partial S_t^2}$$

Then:

$$\begin{aligned}
\mathcal{L}_{BS}(\sigma(y))R(t, S_t, Y_t) &= -S_t\sigma(y_t)\rho_t \frac{\partial^2 P_{0,0}}{\partial S \partial y} \\
&+ (T-t)\left(S_t\sigma(y)\frac{\partial^2}{\partial S\partial y}\mathcal{L}_{BS}(\sigma(y))\, P_{0,0} - \rho_t\sigma^2(y)\sigma'(y)\mathcal{D}_{1,S}\mathcal{D}_{2,S}P_{0,0}\right) \\
&+ (T-t)\sigma^2(y_t)\sigma'(y_t)\rho_t\mathcal{D}_{1,S}\mathcal{D}_{2,S}P_{0,0} \\
&- \frac{(T-t)^2}{2}\sigma^2(y_t)\sigma'(y_t)\rho_t\mathcal{D}_{1,S}\mathcal{D}_{2,S}\mathcal{L}_{BS}(\sigma(y))\, P_{0,0}.
\end{aligned}$$

Since $\mathcal{L}_{BS}(\sigma(y))\, P_{0,0} = 0$, it can be deduced that:

$$\mathcal{L}_{BS}(\sigma(y))R(t, S_t, y_t) \;\; = \;\; -S_t\sigma(y_t)\rho_t \frac{\partial^2 P_{0,0}}{\partial S \partial y}$$

and $R(T, S_T, y_T) = 0$. Then, it follows that:

$$P_{1,0}(t, S_t, y_t) \;\; = \;\; R(t, S_t, y_t)$$

$$P_{1,0}(t, S_t, Y_t) \;\; = \;\; (T-t)S_t\sigma(y_t)\rho_t\frac{\partial^2 P_{0,0}}{\partial S\partial y} - \frac{(T-t)^2}{2}\sigma^2(y_t)\sigma'(y_t)\rho_t S_t\frac{\partial}{\partial S_t}(S_t^2\frac{\partial^2 P_{0,0}}{\partial S_t^2}).$$

Using the relation between the Gamma and the Vega in the Black-Scholes model, it follows that:

$$\begin{aligned}
P_{1,0}(t, S_t, Y_t) &= (T-t)S_t\sigma(y_t)\sigma'(y_t)\rho_t\frac{\partial^2 P_{0,0}}{\partial S\partial\sigma(y)} - \frac{(T-t)^2}{2}\sigma^2(y_t)\sigma'(y_t)\rho_t\frac{S_t}{\sigma(y)(T-t)}\frac{\partial^2 P_{0,0}}{\partial S\partial\sigma(y)} \\
&= \frac{T-t}{2}S_t\sigma(y_t)\sigma'(y_t)\rho_t\frac{\partial^2 P_{0,0}}{\partial S\partial\sigma(y)}
\end{aligned}$$

Finally, $P_{0,1}$ satisfies the equation in the $(0,1)$-term (of order 1 in λ and 0 in b):

$$\mathcal{L}_{BS}(\sigma(y))\, P_{0,1} \;\; = \;\; S_t k\frac{\partial P_{0,0}}{\partial S} - \phi(P_{0,0}),$$

with the final condition:

$$P_{0,1}(T, S_T, Y_T) \;\; = \;\; 0.$$

Then, since the operator $\mathcal{L}_{BS}(\sigma(y))$ commits with the operators ϕ and $S\frac{\partial}{\partial S}$, it follows:

$$P_{0,1}(t, S_t, Y_t) \;\; = \;\; (T-t)(\phi(P_{0,0}) - \frac{\partial P_{0,0}}{\partial S}S_t k).$$

As a conclusion, the option price P can be approximated at order 1 in b and λ by $\hat{P}$:

$$\hat{P}(t, S_t, y_t) = P_{0,0} + bP_{1,0} + \lambda P_{0,1},$$

where:

$$\begin{aligned} P_{0,0}(t, S_t, y_t) &= P_{BS}(t, S_t, \sigma(y_t), K, T), \\ P_{1,0}(t, S_t, y_t) &= \frac{T-t}{2} S_t \sigma(y_t)\sigma'(y_t)\rho_t \frac{\partial^2 P_{0,0}}{\partial S \partial \sigma}, \\ P_{0,1}(t, S_t, y_t) &= (T-t)(\phi(P_{0,0}) - S_t k \frac{\partial P_{0,0}}{\partial S}). \end{aligned}$$

6.3 Appendix 3

The option price $P^{K,T}$ can be approximated at order 1 in b and λ as:

$$P^{K,T} = P_{0,0} + bP_{1,0} + \lambda P_{0,1} + O\big(b^2 + \lambda^2 + b\lambda\big).$$

Let $I_t(K,T)$ be the implied volatility of $P^{K,T}$, which means that $P_t^{K,T} = P_{BS}(t, S_t, I_t(K,T), K, T)$. Recall that if $b = \lambda = 0$, then the constant volatility model is recovered and $I_t(K,T) = \sigma(y_t)$, therefore $I_t(K,T)$ can be written at first order in b and λ as:

$$I_t(K,T) = \sigma(y) + bI_{1,t}(K,T) + \lambda I_{2,t}(K,T) + O\big(b^2 + \lambda^2 + b\lambda\big)$$

Thus, we can perform the following Taylor development:

$$P_t^{K,T} = P_{BS}(t, S_t, \sigma(y_t)) + \frac{\partial P_{BS}}{\partial \sigma}_{|\sigma(y)} (bI_{1,t}(K,T) + \lambda I_{2,t}(K,T)) + O\big(b^2 + \lambda^2 + b\lambda\big)$$

Since $P_{0,0} = P_{BS}(t, S_t, \sigma(y), K, T)$, then by equalizing terms which have the same order in b and λ, it can be deduced that:

$$\begin{aligned} \frac{\partial P_{BS}}{\partial \sigma}_{|\sigma(y_t)} I_{1,t}(K,T) &= P_{1,0} = \frac{T-t}{2} S_t \sigma(y_t)\sigma'(y_t)\rho_t \frac{\partial^2 P_{0,0}}{\partial S \partial \sigma} \\ \frac{\partial P_{BS}}{\partial \sigma}_{|\sigma(y_t)} I_{2,t}(K,T) &= P_{0,1} = (T-t)(\phi(P_{0,0}) - kS_t \frac{\partial P_{0,0}}{\partial S}) \end{aligned}$$

The term $P_{0,0}$ is a Black-Scholes price, then $\frac{\partial^2 P_{0,0}}{\partial S \partial \sigma}$ and $\frac{\partial P_{BS}}{\partial \sigma}$ have the following analytic expressions:

$$\begin{aligned} \frac{\partial^2 P_{0,0}}{\partial S \partial \sigma(y)} &= -\frac{n(d_1)d_2}{\sigma(y)} \\ \frac{\partial P_{0,0}}{\partial \sigma(y)} &= S_t n(d_1)\sqrt{T-t} \end{aligned}$$

where:

$$\begin{aligned} d_1 &= \frac{\log(\frac{F_{t,T}}{K}) + \frac{\sigma(y)^2}{2}(T-t)}{\sigma(y)\sqrt{T-t}} \\ d_2 &= \frac{\log(\frac{F_{t,T}}{K}) - \frac{\sigma(y)^2}{2}(T-t)}{\sigma(y)\sqrt{T-t}} \\ F_{t,T} &= S_t e^{r(T-t)} \end{aligned}$$

By doing necessary calculations, it can be deduced that:

$$I_{1,t}(K,T) \quad = \quad \frac{\rho_t}{2}\frac{\sigma'(y)}{\sigma(y)}\log(\frac{K}{F_{t,T}}) + \rho_t\sigma'(y)\frac{\sigma(y)}{4}(T-t)$$

We aim now to give the expression of $I_2(K,T)$. Recall that:

$$P_{0,1}(t,S_t,y_t) \quad = \quad (T-t)\left(\phi(P_{0,0}) - kS_t\frac{\partial P_{0,0}}{\partial S}\right)$$

Conditional on the variable J, and under the hypothesis that $(J-1)\sim 0$, the quantity $P_{0,0}(S_tJ)$ can be written as:

$$P_{0,0}(S_tJ) = P_{0,0}(S_t) + S_t(J-1)\frac{\partial P_{0,0}}{\partial S}(S_t) + \frac{1}{2}S_t^2(J-1)^2\frac{\partial^2 P_{0,0}}{\partial S^2}(S_t) + \frac{1}{6}S_t^3(J-1)^3\frac{\partial^3 P_{0,0}}{\partial S^3}(S_t) + \frac{1}{24}S_t^4(J-1)^4\frac{\partial^4 P_{0,0}}{\partial S^4}(S_t) + O\big((J-1)^5\big)$$

Then, we can take the expectation of the previous equation with respect to the distribution of J, and we deduce that:

$$E(P_{0,0}(S_tJ)) = P_{0,0}(S_t) + S_tk\frac{\partial P_{0,0}}{\partial S}(S_t) + \frac{1}{2}S_t^2E((J-1)^2)\frac{\partial^2 P_{0,0}}{\partial S^2}(S_t) + \frac{1}{6}S_t^3E((J-1)^3)\frac{\partial^3 P_{0,0}}{\partial S^3}(S_t) + \frac{1}{24}S_t^4E((J-1)^4)\frac{\partial^4 P_{0,0}}{\partial S^4}(S_t) + O\big(E((J-1)^5)\big)$$

We can then write that:

$$I_{2,t}(K,T) = (T-t)\frac{\frac{1}{2}\frac{\partial^2 P_{0,0}}{\partial S^2}(S_t)S_t^2E((J-1)^2) + \frac{1}{6}\frac{\partial^3 P_{0,0}}{\partial S^3}(S_t)S_t^3E((J-1)^3) + \frac{1}{24}\frac{\partial^4 P_{0,0}}{\partial S^4}(S_t)S_t^4E((J-1)^4)}{\frac{\partial P_{BS}}{\partial\sigma(y)}} + O\big(E((J-1)^5)\big)$$

Here again, since $P_{0,0}$ is a Black-Scholes option price, the quantities $\frac{\partial^3 P_{0,0}}{\partial S^3}$ and $\frac{\partial^4 P_{0,0}}{\partial S^4}$ have analytic expressions:

$$\frac{\partial^3 P_{0,0}}{\partial S^3}(S_t) \quad = \quad -\frac{n(d_1)}{\sigma(y)\sqrt{T-t}S_t^2}(1+\frac{d_1}{\sigma(y)\sqrt{T-t}})$$

and:

$$\frac{\partial^4 P_{0,0}}{\partial S^4}(S_t) = \frac{1}{\sigma(y)\sqrt{T-t}}(\frac{d_1n(d_1)}{S_t^2}\frac{1}{S_t\sigma(y)\sqrt{T-t}} + \frac{2n(d_1)}{S_t^3}) - \frac{1}{\sigma^2(y)(T-t)}(\frac{(1-d_1^2)n(d_1)}{S_t^2}\frac{1}{S_t\sigma(y)\sqrt{T-t}} - \frac{2d_1n(d_1)}{S_t^3})$$

Then using the definition of d_1, it can be deduced that:

$$I_{2,t}(K,T) = \frac{E((J-1)^2)}{2\sigma(y)} - \frac{E((J-1)^3)}{4\sigma(y)} + \frac{E((J-1)^3)}{6\sigma^3(y)(T-t)}\log(\frac{K}{F_{t,T}}) + \frac{E((J-1)^4)}{24}(-\frac{4\log(\frac{K}{F_t})}{\sigma^3(y)(T-t)} + \frac{\log(\frac{K}{F_t})^2}{\sigma^5(y)(T-t)^2} + \frac{15}{4\sigma(y)} - \frac{1}{\sigma^3(y)(T-t)}) + O\big(E((J-1)^5)\big)$$

In conclusion, the implied volatility $I_t(K,T)$ writes $I_t(K,T) = \hat{I}_t(K,T) + O(b^2 + \lambda^2 + b\lambda)$ where:

$$\hat{I}_t(K,T) \;=\; d(\rho,b,\lambda,T-t) + a(\rho,b,\lambda,T-t)\log(\frac{K}{F_{t,T}}) + c(\rho,b,\lambda,T-t)\log(\frac{K}{F_{t,T}})^2$$

and the functions a, c and d are summarized as:

$$\begin{aligned} a(\rho,b,\lambda,T-t) &= \frac{b_t\rho_t}{2}\frac{\sigma'(y)}{\sigma(y)} + \lambda_t\frac{E((J-1)^3)}{6\sigma^3(y)(T-t)} - \frac{4\lambda_t E((J-1)^4)}{24\sigma^3(y)(T-t)}, \\ c(\rho,b,\lambda,T-t) &= \frac{\lambda_t E((J-1)^4)}{24\sigma^5(y)(T-t)^2}. \end{aligned}$$

$$\begin{aligned} d(\rho,b,\lambda,T-t) = \sigma(y) + b_t\rho_t\sigma'(y)\frac{\sigma(y)}{4}(T-t) + \lambda_t\frac{E((J-1)^2)}{2\sigma(y)} - \lambda_t\frac{E((J-1)^3)}{4\sigma(y)} \\ + \frac{\lambda_t}{24}E((J-1)^4)(\frac{15}{4\sigma(y)} - \frac{1}{\sigma^3(y)(T-t)}) \end{aligned}$$

6.4 Appendix 4

At order 0 in x_L, the hedging ratio n_H writes:

$$n_H \;=\; \frac{n(d_{1,L})}{n(d_{1,H})}$$

then:

$$\begin{aligned} \Gamma_t &= \frac{1}{2}S_t^2\sigma_{BS}^2(K_L,T)\left(\frac{\partial^2 P_{BS}^L}{\partial S^2} - \frac{\sigma_{BS}^2(K_H,T)}{\sigma_{BS}^2(K_L,T)}\frac{n(d_{1,L})}{n(d_{1,H})}\frac{\partial^2 P_{BS}^H}{\partial S^2}\right), \\ &= \frac{1}{2}S_t^2\sigma_{BS}^2(K_L,T)\left(\frac{\partial^2 P_{BS}^L}{\partial S^2} - \frac{\sigma_{BS}(K_H,T)}{\sigma_{BS}(K_L,T)}\frac{\frac{\partial^2 P_{BS}^L}{\partial S^2}}{\frac{\partial^2 P_{BS}^H}{\partial S^2}}\frac{\partial^2 P_{BS}^H}{\partial S^2}\right), \\ &= \frac{1}{2}S_t^2\sigma_{BS}^2(K_L,T)\left(\frac{\partial^2 P_{BS}^L}{\partial S^2} - \frac{\sigma_{BS}(K_H,T)}{\sigma_{BS}(K_L,T)}\frac{\partial^2 P_{BS}^L}{\partial S^2}\right), \\ &= \frac{1}{2}S_t^2\sigma_{BS}(K_L,T)\frac{\partial^2 P_{BS}^L}{\partial S^2}\left(\sigma_{BS}(K_L,T) - \sigma_{BS}(K_H,T)\right), \end{aligned}$$

Thus, the discriminant D_t of the quadratic equation can be approximated at order 2 in x_L as following:

$$\begin{aligned} D_t &= \left(\frac{n(d_1^L)x_L}{\sigma_{ATM}^2\sqrt{\tau}}\right)^2 S_t\mathcal{S}_T^2 d\left\langle\log(S)\right\rangle_t^2 + 2\Gamma_t dt\left(\frac{S_t n(d_1^L)}{\sigma_{ATM}^3\sqrt{\tau}}x_L^2\right)\mathcal{S}_T^2 d\left\langle\log(S)\right\rangle_t, \\ &= \frac{n(d_1^L)^2 x_L^2 S_t^2\mathcal{S}_T^2}{\tau\sigma_{ATM}^3}d\left\langle\log(S)\right\rangle_t dt\left(\sigma_{BS}(K_L,T) - \sigma_{BS}(K_H,T) + \frac{d\left\langle\log(S)\right\rangle_t}{\sigma_{ATM}(T)dt}\right). \end{aligned}$$

For $T \sim 0$, we can make the approximation $d\left\langle\log(S)\right\rangle_t = \sigma_{ATM}^2(T)dt$ and then $D_t \geq 0$ if $\sigma_{BS}(K_L,T) - \sigma_{BS}(K_H,T) + \sigma_{ATM}(T) \geq 0$.

References

[Bekaert and Wu(2000)] Geert Bekaert and Guojun Wu. Asymmetric volatility and risk in equity markets. *Review of Financial Studies*, 13(1):1–42, 2000.

[Bergomi(2009)] Lorenzo Bergomi. Smile dynamics iv. *Risk*, 12:94–100, 2009.

[Bollerslev et al.(2006)Bollerslev, Litvinova, and Tauchen] Tim Bollerslev, Julia Litvinova, and George Tauchen. Leverage and volatility feedback effects in high-frequency data. *Journal of Financial Econometrics*, 4(3):353–384, 2006.

[Bouchaud et al.(2001)Bouchaud, Matacz, and Potters] Jean-Philippe Bouchaud, Andrew Matacz, and Marc Potters. Leverage effect in financial markets: The retarded volatility model. *Physical Review Letters*, 87(22):228701, 2001.

[Ciliberti et al.(2009)Ciliberti, Bouchaud, and Potters] Stefano Ciliberti, Jean-Philippe Bouchaud, and Marc Potters. Smile dynamics: A theory of the implied leverage effect. *Wilmott Journal*, 1(2):87–94, 2009.

[Vargas et al.(2013)Vargas, Dao, and Bouchaud] Vincent Vargas, Tung-Lam Dao, and Jean-Philippe Bouchaud. Skew and implied leverage effect: smile dynamics revisited. *arXiv preprint arXiv:1311.4078*, 2013.