Linking the performance of vanilla options to the volatility premium

Olivier Daviaud * Abhishek Mukhopadhyay†
Société Générale Société Générale

February 18, 2022

Abstract

The advent of quantitative investing has made it increasingly important to understand the performance drivers of systematic strategies that use derivatives, such as those based on the sale of options. In this paper we introduce a new formulaic representation to analyse the performance of delta-hedged vanilla options, and as a by product provide an explicit link between that performance and the so called volatility premium. We then apply that formula to break down and analyse the historical performance of foreign exchange straddles held to maturity. These results can be used to size option strategies and to disentangle the impact of the various performance drivers in the booming field of systematic options strategies.

1 Introduction

1.1 Motivations

Options pricing has traditionally been approached from the angle of the market maker or the risk manager. As a result, it focuses on calculating

*The authors would like to thank Sandrine Ungari for reading several drafts and offering helpful comments.

†**Disclaimer**: this document was written by Olivier Daviaud and Abhishek Mukhopadhyay at Société Générale. Opinions expressed are current as of the date appearing in this publication and neither Société Générale nor its subsidiaries or affiliates accept any responsibility for liability arising from the use of all or any part of this document.

prices or instantaneous hedging ratios. By contrast, studies on the statistical properties of derivatives held over longer periods of time are less common. At the same time, the advent of quantitative investing has made it increasingly important to understand the performance drivers of systematic strategies which use derivatives, such as those based on the sale of options. The issue is all the more pressing since without a clear statistical framework, the science of strategy selection tends to rely excessively on backtests, as argued in (Lopez De Prado, 2018).

Consider one of the most canonical examples of this kind of strategy: the systematic sale of straddles, i.e. of a call and a put, both struck at the money. It is commonly accepted that the cumulative returns from this kind of strategy arise from the difference between ex-ante implied volatility and ex-post realised volatility, a quantity known as the *volatility premium*. That intuition builds from the Black-Scholes formula, which states that when implied volatility is constant, the P&L of a delta-hedged options portfolio over a short time interval is proportional to gamma times implied minus realised volatility. But that heuristic is insufficient to explain the relationship between the option's performance and the volatility premium when the holding period is of arbitrary length, and when volatility isn't constant.

In particular it gives us no insight into sizing. Let's assume for example that we have a view on what the volatility premium will be in the future for a given asset. Given that prior, what amount of straddle should I sell and delta hedge to reach a given profit target at expiry? It does not explain either why performance and volatility premium are sometimes disconnected. Selling delta hedged straddles on 1y AUDUSD options over the past 13 years and rolling them at expiry would have generated a significant profit for example, in spite of implied volatility trading below realised over that period. We lack a tool, too, to understand the impact of the variations in implied volatility on the performance at expiry of delta hedged options. Is it small and mostly contributing to noise through the option vega, as some practitioners are inclined to think, or can it play a significant role? Finally, what is the ex-ante expected gain or loss in a stationary scenario where the relevant performance drivers do not move, a quantity typically known to practitioners as carry? All these questions require a formula linking the ex-post performance of a delta hedged option to its drivers.

1.2 Volatility premium and option performance explanation: the state of the field

The study of option performance is at the crossroads of two areas of research: risk premium in the options market on the one hand, and pricing theory for risk management and P&L explanation purposes on the other.

On average, the volatility premium is positive: implied volatility tends to be higher than realised volatility, as observed in (Carr & Wu, 2009) for example. Options can be thought of as insurance contracts, and option buyers are willing to pay a premium versus fair value in exchange for protection in an adverse market environment. That source of risk premium has been the subject of many studies over the years, both from a modelling and from an empirical standpoint. (Bakshi & Kapadia, 2003) and later (Hu & Jacobs, 2020) found the volatility premium to be positively correlated with the level of volatility, and in (Carr & Wu, 2015), the authors introduced a model of implied volatility able to match the surface of option prices while allowing for a non-zero premium. Recent empirical studies include (Fritzsch, Irresberger, & Weiß, 2021), which uses machine learning techniques to flag outliers in the relationship between implied and realised volatility and monetise these dislocations. The topic is also of keen interest to banking and asset management practitioners. (Société Générale Cross-Asset Quant Research, 2017, p180), for example, explores the possibility that the level of volatility premium correlates with future options returns.

Several tools have been developed to analyse the performance of vanilla options and to understand the impact and interplay between the underlying drivers. (Ahmad & Wilmott, 2005) measures the impact of the choice of delta hedging methodology, and more specifically of the volatility parameter in the delta calculation. The end goal there is to choose a measure of volatility that will maximise performance. By contrast, our aim is to analyse performance when holding an option over an arbitrary period of time, delta hedging using the live implied volatility. (Bergomi, 2016, p 180) studies the interplay between the changes in value that arise from changes in implied volatility and those that are due to the passage of time. Here the author is interested in timer options, i.e. options which expire once a given target of realised volatility has been reached. Even though his aim is different from ours, some of the arithmetic involved in his calculations also appear in our study.

To the best of our knowledge, our results are novel, both from a theoretical and an empirical standpoint. They do not appear in the literature mentioned above, or in specialised surveys like (James, Fullwod, & Billington, 2015).

1.3 Main results

Starting with the generic P&L formula for a delta hedged option portfolio over an infinitesimal amount of time, we rewrite the equation to disentangle the transfers between the various P&L blocks and integrate it since trade inception. Several terms of interest appear, including one which is proportional to the volatility premium, defined as traded minus realised volatility; another which is equal to the trade's vega times the change in implied volatility; and a third one which is linked to the covariance between gamma and the instantaneous realised variance.

As an immediate application, practitioners can use this formula to calculate the volatility premium's performance contribution ex-post. As we shall see, they can also use it to measure the performance sensitivity to the volatility premium ex-ante, and therefore to size their trade.

In the second part of the paper we use straddles on foreign exchange to illustrate that formula. We find that the volatility premium is the largest performance contributor on average, and that over time that performance component is reminiscent of a volatility swap. In particular this observation validates the choice of selling delta hedged vanilla options systematically as a way to monetise the volatility premium.

We will now present our mathematical results, and use them to analyse the performance of short straddle strategies on G10 currency pairs over the past 13 years.

2 An explicit link between option P&L and volatility premium

Let $P(S, K, \sigma, r, q, t, T)$ be the Black-Scholes price at time t of an option with expiry T, strike K where r the risk-free rate, q the underlying's dividend rate, σ the implied volatility and S the underlying's price. Throughout this paper we use the Black-Scholes formula, and therefore implied volatility, as the value representation for the option. In doing so we align ourselves with (Carr & Wu, 2015) and (Bergomi, 2016), among others. This choice has no consequence on the dynamics of the option prices themselves, but implicitly determines the identity of the performance drivers which appear in the P&L attribution. We shall also assume throughout that the option is delta hedged using the Black-Scholes delta . While from a hedger's perspective the delta scheme is dictated by the dynamics of the underlying (see (Bergomi, 2004) or (Schweizer, 1999)), from an investor's standpoint the choice of delta is a feature of the trading strategy whose optimisation falls beyond the scope of our study. The Black-Scholes delta has the merit of being the simplest hedging scheme of all, and of being used by a number of participants and investable strategies which makes it a natural starting point for our attribution analysis.

If we assume for the time being that the strike-K implied volatility σ is fixed, and that S is an arbitrary continuous process (we shall formalise this in a minute), then the performance of that delta hedged option over a small time window is:

$$\begin{aligned}
P\&L_{[t,t+dt]} &= d\left(e^{-rt}P\right) - \frac{\partial P}{\partial S}e^{-qt}d\left(e^{-(r-q)t}S\right) \\
&= e^{-rt}\left[-rPdt + \frac{\partial P}{\partial t}dt + \frac{\partial P}{\partial S}dS + \frac{1}{2}\frac{\partial^2 P}{\partial S^2}d\left\langle S\right\rangle\right] \\
&\quad - e^{-rt}\frac{\partial P}{\partial S}\left[-(r-q)Sdt + dS\right] \\
&= e^{-rt}\left[-rPdt + \frac{\partial P}{\partial t}dt + \frac{1}{2}\frac{\partial^2 P}{\partial S^2}d\left\langle S\right\rangle + (r-q)S\frac{\partial P}{\partial S}dt\right], \qquad (1)
\end{aligned}$$

where $\langle . \rangle$ denotes quadratic variations. By design, P satisfies the Black-Scholes equation:

$$-rPdt + \frac{\partial P}{\partial t}dt + \frac{1}{2}\frac{\partial^2 P}{\partial S^2}\sigma^2 S^2 dt + (r-q)S\frac{\partial P}{\partial S}dt = 0$$

Therefore (1) becomes

$$P\&L_{[t,t+dt]} = e^{-rt}\left[\frac{1}{2}\frac{\partial^2 P}{\partial S^2}\left(d\left\langle S\right\rangle - \sigma^2 S^2 dt\right)\right] \qquad (2)$$

Our goal is to generalise this equation to non-constant implied volatility and to arbitrary time intervals with an eye to disentangling the various P&L drivers and to making the volatility premium appear.

We first release the assumption that implied volatility is constant and instead let (σ, S) follow an Itô process (see (Revuz & Yor, 1998, p296)) under the physical measure $\mathbb{P}$ and filtration $\mathcal{F}_t$. Our intent here is to make our results as model free as possible, while acknowledging at the same time that Itô calculus operates globally rather than at the path level, and requires a probability measure, i.e. a model. The class of Itô processes is broad and includes all stochastic volatility models. Its main restriction is that it excludes jump diffusion processes, for which (1) would require an extra term (see (Bergomi, 2016, p409)). Finally, we assume for simplicity that interest rates are constant.

We have

$$P\&L_{[t,t+dt]} = e^{-rt}\left[\frac{1}{2}\frac{\partial^2 P}{\partial S^2}\left(d\langle S\rangle - \sigma^2 S^2 dt\right) + \frac{\partial P}{\partial \sigma}d\sigma \right. \\ \left. + \frac{1}{2}\frac{\partial^2 P}{\partial \sigma^2}d\langle\sigma\rangle + \frac{\partial^2 P}{\partial\sigma\partial S}d\langle S,\sigma\rangle\right] \quad (3)$$

One can see from this formula that P&L transfers are likely to occur between the gamma and the vega block. Indeed assume that one is long the vanilla option and therefore that both gamma and vega are positive. If implied volatility suddenly rises, then the vega term records a gain. But because σ is now higher, the bar to clear for the gamma term to make money rises too, meaning that future P&L prospects for that block have worsened.

We pursue two goals: disentangling those two terms so as to have non-overlapping sources of $P\&L$, and forcing the volatility premium to appear into the equation by injecting the volatility traded at inception, σ_0. Once that is done we integrate the resulting formula between trade inception ($t = 0$) and an arbitrary time t. The details can be found in Appendix A, and lead to

$$P\&L_{[0,t]} = \Bigg[\overbrace{\left(t\overline{\Gamma^*} \frac{\sigma_t^r + \sigma_0}{2} \right) (\sigma_t^r - \sigma_0)}^{\text{Volatility premium component}} + \overbrace{\frac{t}{2} \rho \ \mathrm{Stdev}(\Gamma^*) \ \mathrm{Stdev} \left(\frac{d \langle S \rangle}{S^2 du} \right)}^{\text{Gamma covariance effect}}$$

$$+ \underbrace{e^{-rt} \frac{\sigma_t + \sigma_0}{2\sigma_t} \mathcal{V}(t) (\sigma_t - \sigma_0)}_{\text{Vega term}} - \underbrace{\int_0^t \frac{(T-u)}{2} (\sigma_u^2 - \sigma_0^2) d\Gamma^*}_{\text{dGamma term}}$$

$$\underbrace{- \int_0^t \frac{e^{-ru}}{2} \left(\frac{\mathcal{V}}{\sigma} - \frac{\partial \mathcal{V}}{\partial \sigma} \right) d \langle \sigma \rangle}_{\text{Residual drift term}} \Bigg] \quad (4)$$

where

- $\Gamma^* := e^{-rt} S^2 \Gamma$ is the discounted dollar gamma
- $\bar{(.)}$ and Stdev(.) are respectively the sample average and sample standard deviation of any function between 0 and t:

$$\bar{f} := \frac{1}{t} \int_0^t f(u) du, \qquad \mathrm{Stdev}(f) := \sqrt{\frac{1}{t} \int_0^t (f(u) - \bar{f})^2 du}. \quad (5)$$

- σ_t is the implied volatility at t and σ_t^r the realised volatility over $(0, t)$:

$$\sigma_t^r := \sqrt{\frac{1}{t} \int_0^t \frac{d \langle S \rangle}{S^2}} \quad (6)$$

- $d \langle S \rangle / du$ is the Radon–Nikodym derivative of $d \langle S \rangle$ with respect to the Lebesgue measure du
- $\mathcal{V} := \partial P / \partial \sigma$ is the option's vega
- ρ is the sample correlation between the discounted dollar gamma and the instantaneous realised variance

$$\rho = \frac{\frac{1}{t} \int_0^t \left(\Gamma^* - \overline{\Gamma^*} \right) \left(\frac{d\langle S \rangle}{S^2 du} - \overline{\frac{d\langle S \rangle}{S^2 du}} \right) du}{\sqrt{\frac{1}{t} \int_0^t \left(\Gamma^* - \overline{\Gamma^*} \right)^2 du} \sqrt{\frac{1}{t} \int_0^t \left(\frac{d\langle S \rangle}{S^2 du} - \overline{\frac{d\langle S \rangle}{S^2 du}} \right)^2 du}}. \quad (7)$$

Let's review each component in turn, using the nomenclature introduced in (4):

- **Volatility premium component**. This term is equal to the volatility premium, $\sigma_t^r - \sigma_0$, times a scaling factor, which we will call the *volatility premium scaling factor* and which is equal to $t\overline{\Gamma^*}(\sigma_t^r + \sigma_0)/2$. It measures the volatility premium's contribution to performance ex-post. It can also be used ex-ante to estimate the expected performance sensitivity to the volatility premium. Indeed Γ^* is a martingale: to prove it, observe that since P is homogeneous as a function of S and K, Euler's homogeneous function theorem entails that $S^2\partial^2 P/\partial S^2 = K^2\partial^2 P/\partial K^2$. Now the right hand side is a martingale because the second derivative of P with respect to K is the discounted expectation of a Dirac around the strike at maturity (see (Bergomi, 2016)[page 29]). Under these circumstances the expected value of $\overline{\Gamma^*}$ is equal to Γ^* at $t = 0$, hereafter written Γ_0^*. Assuming that σ_t^r is known at inception, or equivalently that we have a view on where volatility will realise between 0 and t, and leaving aside the impact of correlation between σ^r and $\overline{\Gamma^*}$, the time-0, risk neutral expectation of the volatility premium scaling factor $t\overline{\Gamma^*}(\sigma_t^r + \sigma_0)/2$ can be approximated by $t\Gamma_0^*(\sigma_t^r + \sigma_0)/2$. From (8) this is equal to $t(\mathcal{V}_0/\sigma_0 T)(\sigma_t^r + \sigma_0)/2$, or $(t/T)\mathcal{V}_0(1 + (\sigma_t^r - \sigma_0)/2\sigma_0)$ after rearranging the σ_0's.

- **Gamma covariance effect.** This term is equal to the sample covariance between the discounted dollar gamma Γ^* and the instantaneous realised variance, $d\langle S\rangle/S^2 du$. It can be thought of as a corrective term. As we shall see in the next section, its contribution can be significant.

- **Vega term.** This term is proportional to the discounted vega at time t times $\sigma(t) - \sigma_0$, and can be interpreted as the fraction of performance which comes from directional changes in implied volatility. In particular it is zero at expiry, as vega vanishes there. This shows that when an option is held to maturity, directional moves in implied volatility do not impact terminal P&L directly. The only impact occurs through changes in exposure to other factors, such as the volatility premium scaling factor.

- **dGamma term** As we shall see empirically, the total contribution from this term is often small in our backtest. The martingale property of Γ^*, alluded to above, is likely at play here since it guarantees that this integral has zero expectation under the risk neutral measure. Note however that our backtest takes place under the physical measure, and that the expectation under that measure need not be exactly zero[1].

- **Residual drift terms** This term is linked to volatility of volatility, since this is one of the drivers of the quadratic variation of σ. For straddles, we shall see that its total contribution is rather small, on average.

3 Empirical results

In this section, we use the results from the above section to analyse the performance of a well-known option strategy: selling a straddle, hedging it at regular intervals and holding it to maturity. We focus on foreign exchange underlyings.

Our backtester relies on end-of-day data for implied volatility, and end-of-day and 10am NY fixing data for FX spot. It rebalances the delta at the London close every business day and calculates it based on the spot level observed at the 10am NY fixing and on the volatility level observed on the prior day's close. This lag is introduced to make the backtest as realistic as possible. We focus on some of the main G7 currency pairs, namely EURUSD, USDJPY, GBPUSD, EURGBP, USDCAD, AUDUSD, USDNOK, USDSEK, USDCHF and EURCHF, and on one week, one month, six months and 1y option tenors.

3.1 The volatility premium scaling factor

Before we measure and compare the various components of that equation empirically, we need to scale the straddles so as to make the comparison possible across tenors and underlyings. One natural candidate for this measure is the volatility premium scaling factor (cf section 2, page 8). It links the performance of the trade to one of its more intuitive drivers, namely the volatility premium. It is also decorrelated from that quantity, on average, as

[1]We thank Liuren Wu for this observation on the risk neutral measure, and Peter Carr for pointing out to us that Γ^* is a martingale in a broader context than Black-Scholes.

table 1 shows. This decorrelation means that on average the volatility premium component (cf (4)) is equal to the volatility premium times a constant, itself equal to the empirical average of the volatility premium scaling factor.

	AUDUSD	EURCHF	EURGBP	EURUSD	GBPUSD	USDCAD	USDCHF	USDJPY	USDNOK	USDSEK	Avg
1 month	0.05	-0.57	0.17	0.22	0.11	0.09	0.03	0.01	-0.07	-0.01	0.00
1 week	-0.08	-0.75	-0.01	0.03	-0.02	-0.06	-0.10	-0.09	-0.11	-0.11	-0.13
1 year	0.21	-0.10	0.42	0.24	0.30	0.05	0.03	0.22	0.01	0.24	0.16
6 months	0.22	-0.22	0.29	0.21	0.28	0.16	-0.08	0.17	-0.06	0.16	0.11
Avg	0.10	-0.41	0.22	0.18	0.17	0.06	-0.03	0.08	-0.06	0.07	0.04

Table 1: Average correlation between the vol premium scaling factor and the vol premium

As we noted in section 2, the expected value of the volatility premium scaling factor can be approximated by the option's vega at inception times a multiplier which is typically relatively close to 1. On first order, scaling by the average volatility premium scaling factor is therefore similar to scaling by the average inception vega, making the process intuitive. For \$100 invested at every option roll, table 2 shows what the average volatility premium scaling factor looks like, and in the numerical results that follow, all strategies are sized so that their historical volatility premium scaling factor averages \$100.

Tenor	AUDUSD	EURCHF	EURGBP	EURUSD	GBPUSD	USDCAD	USDCHF	USDJPY	USDNOK	USDSEK
1 week	5.39	5.72	5.37	5.35	5.38	5.30	5.40	5.31	5.38	5.30
1 month	2.72	2.80	2.61	2.62	2.69	2.68	2.64	2.78	2.71	2.66
6 months	1.18	1.28	1.11	1.06	1.11	1.19	1.21	1.17	1.15	1.14
1 year	0.81	0.94	0.80	0.75	0.84	0.86	0.94	0.86	0.85	0.79

Table 2: Average volatility scaling factor per year, for \$100 straddle notional.

3.2 Attribution summary

Figure 1 shows what this scaled contribution looks like on average for G10 currency pairs over the past 13 years. We excluded CHF from this universe as it wasn't a free floating currency from 6 September 2011 to 15 January 2015. The chart shows the average yearly backtest P&L, as well as the contribution from the five component in (4). It also overlays the average volatility premium through the life of the backtest ("Vol. premium"), and shows the difference between the backtest P&L and the sum of the five components ("Residual P&L"). This term is non-zero for a few reasons: (4) assumes continuous delta hedging whereas our implementation hedges daily; it assumes constant interest rates whereas our backtest uses market rates; and it precludes jumps.

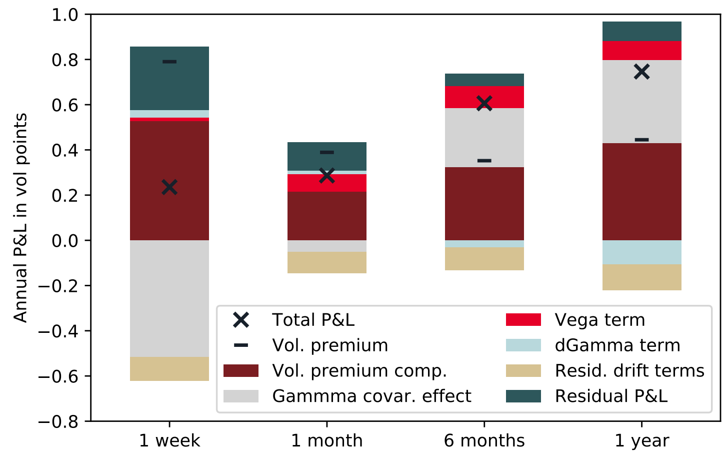

Figure 1: Contribution of each P&L driver to total performance, average across currency pairs excluding CHF

We can draw a number of conclusions from these charts.

The first one is that the volatility premium component of the performance, in brown, is very close to the ex-post volatility premium (dashed markers). This confirms our heuristic from section 3.1, namely that the lack of correlation between the ex post volatility premium and the volatility premium scaling factor made the volatility premium component behave on average as if the volatility premium scaling factor were roughly constant. Put differently, straddles have a performance component which behaves similarly to a volatility swap.

The second conclusion is that the gamma covariance effect, in grey, is the second largest contributor. In particular it solves the AUDUSD paradox laid out in section 1.1, namely that in certain instances short straddle strategies make money even though the volatility premium is negative: in those cases P&L is driven by the gamma covariance effect. Furthermore, both charts 1 and table 3 show that the gamma covariance effect's contribution tends to be negative for short-dated maturities and positive for long-dated ones. Note also that the correlation in table 3 is relatively insensitive to the choice of underlying, an intriguing pattern that calls for further mathematical investigation.

Third, note that the contribution from the vega term is small but not zero, in spite of that term vanishing at expiry according to (4). This is the result of a constraint in our code, which in a few instances forced the backtest to roll the option a few days before expiry.

Fourth, the dGamma term is relatively small, in coherence with the martingale property of Γ^*, as explained in section 2.

Fifth, the residual drift terms are small too.

CurrencyPairs	AUDUSD	EURCHF	EURGBP	EURUSD	GBPUSD	USDCAD	USDCHF	USDJPY	USDNOK	USDSEK	Avg
1 month	-0.04	-0.01	-0.03	-0.04	-0.03	-0.05	-0.05	-0.03	-0.03	-0.05	-0.04
1 week	-0.24	-0.18	-0.21	-0.31	-0.26	-0.24	-0.28	-0.22	-0.23	-0.24	-0.24
1 year	0.04	0.08	0.04	0.05	0.04	0.04	0.02	0.05	0.04	0.04	0.04
6 months	0.03	0.08	0.03	0.03	0.02	0.03	0.02	0.03	0.03	0.02	0.03
Avg	-0.06	-0.01	-0.04	-0.07	-0.06	-0.06	-0.07	-0.04	-0.05	-0.06	-0.05

Table 3: Average path-wise correlation between Γ and instantaneous realised variance

4 Conclusion

In this paper we establish a formal link between the P&L of a delta-hedged option and the so-called volatility premium, thereby providing mathematical confirmation of an intuition that is central to the volatility market. In doing so we create a generic P&L attribution formula for delta-hedged options, valid for a broad class of processes. In addition to allowing for the ex-post measurement of the volatility premium's contribution to P&L, our formula has a number of immediate applications for options trading. It provides a guide to the ex-ante sensitivity of terminal P&L to the volatility premium, linking it to vega at inception. It also explains why in certain instances the volatility premium is not the main driver of P&L. And it shows that once adequately normalised, the terminal P&L is relatively insensitive to directional changes in implied volatility. Topics for future research include further analysis of two terms in the formula - the residual drift term and the gamma covariance effect - to gain full insight into ex-ante P&L and a further understanding of carry for delta hedged options.

A Derivation of the P&L equation

A.1 Disentangling gamma and vega

Starting from (3) we recall that for the Black-Scholes formula, gamma and vega are linked (e.g. (Bergomi, 2016, p 180)):

$$\frac{\partial P}{\partial \sigma} = S^2 \frac{\partial^2 P}{\partial S^2} \sigma (T - t) \tag{8}$$

Focusing for now on the gamma and vega part of (3), we set:

$$P\&L_{GV} := e^{-rt} \left[\frac{1}{2} \frac{\partial^2 P}{\partial S^2} \left(d \langle S \rangle - \sigma^2 S^2 dt \right) + \frac{\partial P}{\partial \sigma} d\sigma \right] \tag{9}$$

Using (8) this becomes:

$$P\&L_{GV} = e^{-rt} \left[\frac{1}{2} \frac{\partial^2 P}{\partial S^2} \left(d \langle S \rangle - \sigma^2 S^2 dt \right) + S^2 \frac{\partial^2 P}{\partial S^2} \sigma (T - t) d\sigma \right]$$

Now note that by Itô: $d\sigma^2 = 2\sigma d\sigma + d \langle \sigma \rangle$. Therefore:

$$P\&L_{GV} = e^{-rt} \frac{1}{2} S^2 \frac{\partial^2 P}{\partial S^2} \left[\left(\frac{d \langle S \rangle}{S^2} - \sigma^2 dt \right) + (T - t)(d\sigma^2 - d \langle \sigma \rangle) \right]$$

Setting $\Gamma_{\$} := S^2 \partial^2 P / \partial S^2$ and $\sigma_0 := \sigma(t_0)$ as the implied volatility at trade date, we obtain:

$$\begin{aligned}
P\&L_{GV} &= e^{-rt} \frac{1}{2} \Gamma_{\$} \left[\left(\frac{d \langle S \rangle}{S^2} - \sigma_0^2 dt \right) + (\sigma_0^2 - \sigma^2) dt + (T - t) \left(d\sigma^2 - d \langle \sigma \rangle \right) \right] \\
&= \frac{e^{-rt}}{2} \Gamma_{\$} \left[\left(\frac{d \langle S \rangle}{S^2} - \sigma_0^2 dt \right) + (\sigma_0^2 - \sigma^2) dt \right. \\
&\quad \left. + (T - t) d\sigma^2 - (T - t) d \langle \sigma \rangle \right] \\
&= \frac{e^{-rt}}{2} \Gamma_{\$} \left[\left(\frac{d \langle S \rangle}{S^2} - \sigma_0^2 dt \right) + d \left((T - t)(\sigma^2 - \sigma_0^2) \right) - (T - t) d \langle \sigma \rangle \right] \\
&= \frac{e^{-rt}}{2} \Gamma_{\$} \left(\frac{d \langle S \rangle}{S^2} - \sigma_0^2 dt \right) + \frac{e^{-rt}}{2} \Gamma_{\$} d \left((T - t)(\sigma^2 - \sigma_0^2) \right) \\
&\quad - \frac{e^{-rt}}{2} \Gamma_{\$} (T - t) d \langle \sigma \rangle
\end{aligned} \tag{10}$$

Let $\Gamma^* := e^{-rt}\Gamma$ be the discounted dollar gamma. Per Itô's formula, for any two functions U and V,

$$d(UV) = UdV + VdU + d\langle UV\rangle .$$

Applying this to the middle term in (10) yields

$$\begin{aligned}P\&L_{GV} = \Bigg[&\frac{\Gamma^*}{2}\left(\frac{d\langle S\rangle}{S^2} - \sigma_0^2 dt\right) + \frac{1}{2}d\left(\Gamma^*(T-t)(\sigma^2-\sigma_0^2)\right)\\ &-\frac{(T-t)}{2}(\sigma^2-\sigma_0^2)d\Gamma^* - \frac{(T-t)}{2}d\left\langle\sigma^2,\Gamma^*\right\rangle - \frac{\Gamma^*}{2}(T-t)d\langle\sigma\rangle\Bigg] \end{aligned} \tag{11}$$

Using (8) we can rewrite the inner part of the middle term as

$$\begin{aligned}\frac{1}{2}\left(\Gamma^*(T-t)(\sigma^2-\sigma_0^2)\right) &= \frac{1}{2}e^{-rt}S^2\frac{\partial^2 P}{\partial S^2}(T-t)(\sigma^2-\sigma_0^2)\\ &= \frac{1}{2}e^{-rt}S^2\frac{\partial^2 P}{\partial S^2}(T-t)\frac{\sigma}{\sigma}(\sigma+\sigma_0)(\sigma-\sigma_0)\\ &= e^{-rt}\frac{\partial P}{\partial\sigma}\frac{\sigma+\sigma_0}{2\sigma}(\sigma-\sigma_0)\end{aligned}$$

Plugging this in (11) yields

$$\begin{aligned}P\&L_{GV} = \Bigg[&\frac{\Gamma^*}{2}\left(\frac{d\langle S\rangle}{S^2} - \sigma_0^2 dt\right) + d\left(e^{-rt}\frac{\sigma+\sigma_0}{2\sigma}\frac{\partial P}{\partial\sigma}(\sigma-\sigma_0)\right)\\ &-\frac{(T-t)}{2}(\sigma^2-\sigma_0^2)d\Gamma^* - \frac{(T-t)}{2}d\left\langle\sigma^2,\Gamma^*\right\rangle - \frac{\Gamma^*}{2}(T-t)d\langle\sigma\rangle\Bigg] \end{aligned} \tag{12}$$

A.2 From gamma P&L to volatility premium

Let's now focus on the gamma component of (12) and set $t_0 := 0$. For any t in $(0, T)$ we have:

$$\int_0^t \Gamma^*\left(\frac{d\langle S\rangle}{S^2} - \sigma_0^2 du\right) = \int_0^t \Gamma^*\frac{d\langle S\rangle}{S^2} - \sigma_0^2\int_0^t \Gamma^* du \tag{13}$$

We denote $\bar{(.)}$ the empirical average of any function between 0 and t as in (5) and let $d\langle S\rangle/du$ be the Radon–Nikodym derivative of $d\langle S\rangle$ with respect to the Lebesgue measure du. We can rewrite the first part of (13) as:

$$\int_0^t \Gamma^*\frac{d\langle S\rangle}{S^2} = t\left(\frac{1}{t}\int_0^t \Gamma^*\frac{d\langle S\rangle}{S^2 du}du\right).$$

Note now that for any functions f,g,

$$\frac{1}{t}\int_0^t f(u)g(u)du = \bar{f}\bar{g} + \frac{1}{t}\int_0^t (f(u)-\bar{f})(g(u)-\bar{g})du.$$

Therefore:

$$\begin{aligned}
\int_0^t \Gamma^* \frac{d\langle S\rangle}{S^2} &= t\left(\frac{1}{t}\int_0^t \Gamma^* \frac{d\langle S\rangle}{S^2 du}du\right) \\
&= t\left(\overline{\Gamma^*}\frac{1}{t}\int_0^t \frac{d\langle S\rangle}{S^2 du}du + \frac{1}{t}\int_0^t \left(\Gamma^* - \overline{\Gamma^*}\right)\left(\frac{d\langle S\rangle}{S^2 du} - \overline{\frac{d\langle S\rangle}{S^2 du}}\right)du\right) \\
&= t\left(\overline{\Gamma^*}\frac{1}{t}\int_0^t \frac{d\langle S\rangle}{S^2}\right. \\
&\qquad \left. +\rho\sqrt{\frac{1}{t}\int_0^t \left(\Gamma^* - \overline{\Gamma^*}\right)^2 du}\sqrt{\frac{1}{t}\int_0^t \left(\frac{d\langle S\rangle}{S^2 du} - \overline{\frac{d\langle S\rangle}{S^2 du}}\right)^2 du}\right)
\end{aligned}$$

where ρ is defined in (7).

Letting σ_t^r be the realised volatility over $(0,t)$ as defined in (6), (13) becomes

$$\int_0^t \Gamma^*\left(\frac{d\langle S\rangle}{S^2} - \sigma_0^2 du\right) = t\overline{\Gamma^*}\left((\sigma_t^r)^2 - \sigma_0^2\right) + t\rho\text{Stdev}(\Gamma^*)\text{Stdev}\left(\frac{d\langle S\rangle}{S^2 du}\right)$$

where Stdev(.) is the sample standard deviation as defined in (5). Using $a^2 - b^2 = (a-b)(a+b)$ and reintroducing the 1/2 factor from (12) we have:

$$\begin{aligned}
\int_0^t \frac{1}{2}\Gamma^*\left(\frac{d\langle S\rangle}{S^2} - \sigma_0^2 du\right) = \left(t\overline{\Gamma^*}\frac{\sigma_t^r + \sigma_0}{2}\right)(\sigma_t^r - \sigma_0) \\
+ \frac{t}{2}\rho\ \text{Stdev}(\Gamma^*)\ \text{Stdev}\left(\frac{d\langle S\rangle}{S^2 du}\right) \quad (14)
\end{aligned}$$

We are now going to merge (3) and (12). Let us first tidy the four quadratic variation terms included in these equations. Letting $\mathcal{V} := \partial P/\partial\sigma$, differentiating (8) gives

$$\sigma(T-t)d\Gamma^* + \Gamma^*(T-t)d\sigma + (T-t)d\langle\sigma,\Gamma^*\rangle - \Gamma^*\sigma dt = e^{-rt}d\mathcal{V} - \mathcal{V}re^{-rt}dt \quad (15)$$

Taking cross variations with σ on both sides and using Itô's formula on $d\mathcal{V}$ to expand $d\langle \mathcal{V}, \sigma \rangle$ gives

$$\sigma(T-t)d\langle \Gamma^*, \sigma \rangle = e^{-rt} \left(\frac{\partial \mathcal{V}}{\partial \sigma} d\langle \sigma \rangle + \frac{\partial \mathcal{V}}{\partial S} d\langle \sigma, S \rangle \right) - \Gamma^*(T-t)d\langle \sigma \rangle \tag{16}$$

After some reshuffling this allows us to rewrite the four quadratic variation blocks in (3) and (12) as

$$\begin{aligned} -\frac{(T-t)}{2} d\left\langle \sigma^2, \Gamma^* \right\rangle - \frac{(T-t)}{2} \Gamma^* d\left\langle \sigma \right\rangle + e^{-rt} \left(\frac{1}{2} \frac{\partial \mathcal{V}}{\partial \sigma} d\left\langle \sigma \right\rangle + \frac{\partial \mathcal{V}}{\partial S} d\left\langle S, \sigma \right\rangle \right) \\ = \frac{e^{-rt}}{2} \left(\frac{\mathcal{V}}{\sigma} - \frac{\partial \mathcal{V}}{\partial \sigma} \right) d\left\langle \sigma \right\rangle \end{aligned} \tag{17}$$

Combining (3), (12), (14) and (17) yields (4).

References

Ahmad, R., & Wilmott, P. (2005). Which free lunch would you like today, sir?: Delta hedging, volatility arbitrage and optimal portfolios. *Wilmott magazine*.

Bakshi, G., & Kapadia, N. (2003, April). Delta-hedged gains and the negative market volatility risk-premium. *The Review of Financial Studies*, *16*(2), 527–566.

Bergomi, L. (2004). Smile dynamics i. *Risk*, *9*, 117–123.

Bergomi, L. (2016). *Stochastic volatility modeling*. Chapman & Hall.

Carr, P., & Wu, L. (2009). Variance risk premiums. *The Review of Financial Studies*, *22*(3), 1311–1341.

Carr, P., & Wu, L. (2015). Analyzing volatility risk and risk premium in option contracts: A new theory. *Journal of Financial Economics*, *120*, 1–20.

Fritzsch, S., Irresberger, F., & Weiß, G. (2021). Cross-section of option returns and the volatility risk premium. *Preprint*.

Hu, G., & Jacobs, K. (2020, May). Volatility and expected option returns. *Journal of Financial and Quantitative Analysis*, *55*(3), 1025–1060.

James, J., Fullwod, J., & Billington, P. (2015). *Fx option performance - an analysis of the value delivered by fx options since the start of the market*. Wiley.

Lopez De Prado, M. (2018). *Advances in financial machine learning*. Wiley.

Revuz, D., & Yor, M. (1998). *Continuous martingales and brownian motion*. Springer.

Schweizer, M. (1999). *A guided tour through quadratic hedging approaches* (Tech. Rep.). SFB 373 Discussion Paper.

Société Générale Cross-Asset Quant Research. (2017, September). *A close-up view of the option premium.*