A Story of Volatility Investing and Trading Part II : The Emergence and Use of Variance Swaps

Study carried out by the Quantitative Practice
Special thanks to Pierre-Edouard THIERY

canopee
empowering ecosystem

www.canopee-group.com

Summary

Introduction

This paper is the second out of a series of three papers entitled "A Story of Volatility Investing and Trading". In our first paper, we insisted on how options can be used to trade volatility. However those approaches, even though they are interesting from a mathematical and academic point of view, are not really efficient when used in real life. Besides all the practical difficulties of optional approaches, the main weakness lies in the fact that such strategies do not provide pure exposure to volatility: investors have to bear other risks than volatility, and those risks may impact their P&L independently from how the volatility has moved.

To circumvent those issues, variance swaps emerged in the 1990s. In this paper, we set forth all the characteristics of variance swaps. Such swaps, besides providing investors with pure volatility exposure, display many practical advantages. For instance, their replication is fairly easy, inasmuch as it is only based on options and futures.

Variance swaps is also a tool which proves to be useful when one tries to make the most of volatility. Volatility is indeed known as one of the most famous alternative risk premia, meaning that it is possible to extract positive returns beyond what the market may offer by investing in volatility. Of course, variance swaps play a significant role to achieve such a result.

1 Variance Swaps, or the Emergence of Pure Volatility Products

Definition 1 (Variance Swap)
A variance swap is simply a forward contract which pays at maturity T the difference between the realized variance of an underlying asset, and a variance strike K_v chosen at the inception of the contract. If $\mathcal{N}$ is the notional of the contract, the payoff at maturity is:

$$(\sigma^2 - K_v)\mathcal{N}$$

where σ^2 is usually computed as the annual variance of the daily log returns of the asset, with the assumption that the mean of those returns is 0:

$$\sigma^2 = \frac{252}{n}\sum_{i=1}^{n} \ln\left(\frac{S_{t_i}}{S_{t_{i-1}}}\right)^2$$

where $n+1$ is the number of days for the period $[0, T]$, and $t_0 = 0 < t_1 < \cdots < t_n = T$.

When an investor is long/short a variance swap, it means he or she receives the realized volatility/variance strike at maturity.

The assumption that the mean of the returns is worth 0 is of practical interest: thanks to it, variance swaps exhibit an additive property, meaning that it is possible to build a variance swap over a longer period of time by considering several variance swaps over smaller, consecutive periods of time. This is also one of the reasons which explain why we prefer to use variance swaps instead of volatility swaps. The proof of this additive property is given in the subsection 1.3, when we introduce the forward starting variance swap.

1.1 Pricing and Replication of Variance Swaps

The interest of variance swaps also lies in the fact that they are easily replicated thanks to extremely common financial instruments: options and futures [1].

Proposition 1 (Replication of a Variance Swap)
A variance swap can be replicated thanks to a continuum of put and call options inversely weighted by the square of their strike, plus a dynamic position in futures. Furthermore the fair forward value of the variance at time $t = 0$ is equal to:

$$\frac{2e^{rT}}{T}\left[\int_0^{F_0}\frac{Put(S_0, 0, K, T)}{K^2}dK + \int_{F_0}^{\infty}\frac{Call(S_0, 0, K, T)}{K^2}dK\right]$$

where F_0 is the futures price of the asset at maturity T from time $t = 0$, i.e. $F_0 = S_0e^{rT}$.

The proof of this proposition is pivotal since it provides us with the replicating strategy of a variance swap. Let us denote F_t the futures price: for any given function f, we can apply the Ito formula to $f(F_t)$:

$$f(F_T) = f(F_0) + \int_0^T f'(F_t)dF_t + \frac{1}{2}\int_0^T F_t^2 f''(F_t)\sigma_t^2 dt$$

where σ_t is the instantaneous volatility of the underlying asset. If we assume that $\frac{dS_t}{S_t} = \mu_t dt + \sigma_t dW_t$, since $F_t = S_t e^{r(T-t)}$, we easily see that the instantaneous volatility of the futures price is the same as the underlying asset.

The idea is to use a function f whose second derivative is equal to $\frac{1}{x^2}$ in order to simplify the expression within the integral. We use:

$$f(x) = \ln\left(\frac{F_0}{x}\right) + \frac{x}{F_0} - 1$$

This gives us:

$$\frac{1}{2}\int_0^T \sigma_t^2 dt = \ln\left(\frac{F_0}{F_T}\right) + \frac{F_T}{F_0} - 1 - \int_0^T\left(\frac{1}{F_0} - \frac{1}{F_t}\right)dF_t$$

This expression is very close to the one of the realized volatility, which is the quantity we have to replicate since this is the quantity which has to be paid at maturity, but which is unknown at the inception of the contract.

The first part of the expression, i.e. $\ln\left(\frac{F_0}{F_T}\right) + \frac{F_T}{F_0} - 1$, can be replicated using the Carr formula. This formula states that, for any function g which is smooth enough, it is possible to replicate an option whose payoff is $g(F_T)$ at maturity using only static positions in call and put options, bonds and futures.

Proposition 2 (Carr Formula)
κ is an arbitrary number, and f a function smooth enough. The terminal payoff $f(F_T)$ can be rewritten:

$$g(F_T) = \underbrace{g(\kappa)}_{A} + \underbrace{g'(\kappa)\left[(F_T-\kappa)^+ - (\kappa - F_T)^+\right]}_{B}$$

$$+ \underbrace{\int_0^{\kappa} g''(\kappa)(K-F_T)^+ dK}_{C} + \underbrace{\int_{\kappa}^{\infty} g''(\kappa)(F_T-K)^+ dK}_{D}$$

Each term in this equation can be interpreted with a financial meaning, thus giving, in the absence of arbitrage, the replication strategy. In order to get A at maturity, all it takes is a position of $g(\kappa)$ zero-coupon bonds paying 1 at maturity T. B corresponds to a position of $g'(\kappa)$ calls with strike κ, and the opposite position of $g'(\kappa)$ puts with strike κ. C is given by a static position of $g''(\kappa)dK$ puts for all the strikes inferior to κ, and D is given by a static position of $g''(\kappa)dK$ calls for all the strikes superior to κ.

In our context, it is useful to take $\kappa = F_0$: both $f(F_0)$ and $f'(F_0)$ are equal to 0. So we just have:

$$\ln\left(\frac{F_0}{F_t}\right) + \frac{F_T}{F_0} - 1 = \int_0^{F_0} \frac{1}{K^2}(K-F_T)^+ dK + \int_{F_0}^{\infty} \frac{1}{K^2}(F_T-K)^+ dK$$

Therefore, in order to replicate $\frac{1}{T}\int_0^T \sigma_t^2 dt$, a trader should buy a continuum of puts with strike prices ranging from 0 to F_0, and a continuum of calls with strike prices ranging from F_0 to infinity, and then roll a position in futures, holding at date t: $2\frac{e^{-r(T-t)}}{T}\left(-\frac{1}{F_0} + \frac{1}{F_t}\right)$.

In order to implement such a strategy, only the options require an initial payment. The initial cost of the strategy replicating $\frac{1}{T}\int_0^T \sigma_t^2 dt$ is:

$$\frac{1}{T}\int_0^{F_0} \frac{2}{K^2} P(S_0,0,T,K)dK + \frac{1}{T}\int_{F_0}^{\infty} \frac{2}{K^2} C(S_0,0,T,K)dK$$

where P and C denote the prices of a put and a call.

The initial cost enables us to compute the fair forward value of the variance at the inception of the contract. This quantity is defined as the level of the variance strike which ensures that the variance swap is fair, i.e. the level of variance the investor paying the realized volatility should get at maturity. Since the investor paying the realized volatility has to pay

$$\frac{1}{T}\int_0^{F_0} \frac{2}{K^2} P(S_0,0,T,K)dK + \frac{1}{T}\int_{F_0}^{\infty} \frac{2}{K^2} C(S_0,0,T,K)dK$$

at $t = 0$ in order to replicate the realized volatility, at maturity he or she should receive an equivalent variance strike. When taking into account the discount factor, the fair value is:

$$K_v^{fair} = \frac{2e^{rT}}{T}\left[\int_0^{F_0} \frac{1}{K^2} P(S_0,0,T,K)dK + \int_{F_0}^{\infty} \frac{1}{K^2} C(S_0,0,T,K)dK\right]$$

1.2 The Counterparty Risk: Capped Variance Swaps

When dealing with variance swaps, it is important to bear in mind that such products are over-the-counter, meaning that each side of the swap faces a counterparty risks: one of the participants may go bankrupt before the end of the contract. This is even more important insofar as being short a variance swap may be extremely hazardous: the investor must pay the realized volatility, and this quantity is not capped, so the investor may have to pay a huge amount of money at maturity. For instance, if the underlying asset defaults – at some time t_i, S_{t_i} is equal to 0 –, the realized volatility becomes infinite. There is no limit to the potential loss one may face when being short a variance swap.

This explains why variance swaps are often capped in real life: the payoff of a capped variance swap [2] is equal to:

$$\left(min(\sigma^2, m \times K_v) - K_v\right)\mathcal{N}$$

where m is the multiplier setting the cap value.

When it comes to replication, capped variance swaps also display a practical advantage. Indeed we saw that, in order to be able to replicate perfectly a variance swap, a continuum of options is required, with strikes up to the infinity. Of course there are no such options in real life, so it is impossible to perfectly hedge a variance swap. On the contrary, a capped variance swap does not require a continuum of options with strikes ranging from 0 to infinity: in real life, implementing the replication of a variance swaps amounts to replicating a capped variance swap.

1.3 The Additive Property of Variance Swaps: Forward Starting Variance Swaps

A forward starting variance swap is a contract struck at a time $t = 0$, but whose payoff is based on what will happen between two future dates T and T': this is a variance swap for the period $\left[T, T'\right]$, negotiated at $t = 0$. Let us assume that the period $\left[0, T'\right]$ is made of $N+1$ days: $t_0 = 0 < t_1 < \cdots < t_n = T < \cdots < t_N = T'$. The payoff of the forward starting variance swap is:

$$\left(\frac{252}{N-n}\sum_{i=n+1}^{N} \ln\left(\frac{S_{t_i}}{S_{t_{i-1}}}\right)^2 - K_v\right)\mathcal{N}$$

Insofar as the payoff of a variance does not take into account the mean of the log returns, which is supposed to be zero, variance swaps exhibit an additive property. To make it simple, using the above notations, a variance swap on $[0, T]$ plus a variance swap on $\left[T, T'\right]$ is equal to a variance swap on $\left[0, T'\right]$. This explains why a forward starting variance swap can be seen as a "calendar spread" of two variance swaps with two different maturity dates: being long a forward starting variance swap amounts to being long a variance swap with a long maturity, and short a variance swap with a short maturity.

Mathematically, we can rewrite the payoff of a forward starting variance swap:

$$\frac{252}{N-n}\sum_{i=n+1}^{N}\ln\left(\frac{S_{t_i}}{S_{t_{i-1}}}\right)^2 - K_v$$

$$= \frac{252}{N-n}\left[\sum_{i=1}^{N}\ln\left(\frac{S_{t_i}}{S_{t_{i-1}}}\right)^2 - \sum_{i=1}^{n}\ln\left(\frac{S_{t_i}}{S_{t_{i-1}}}\right)^2\right] - K_v$$

$$= \frac{252}{N-n}\frac{N}{N}\sum_{i=1}^{N}\ln\left(\frac{S_{t_i}}{S_{t_{i-1}}}\right)^2 - \frac{252}{N-n}\frac{n}{n}\sum_{i=1}^{n}\ln\left(\frac{S_{t_i}}{S_{t_{i-1}}}\right)^2 - K_v$$

$$= \frac{N}{N-n}\frac{252}{N}\sum_{i=1}^{N}\ln\left(\frac{S_{t_i}}{S_{t_{i-1}}}\right)^2 - \frac{n}{N-n}\frac{252}{n}\sum_{i=1}^{n}\ln\left(\frac{S_{t_i}}{S_{t_{i-1}}}\right)^2$$

$$- \left(\frac{N}{N-n} - \frac{n}{N-n}\right)K_v$$

$$= \frac{N}{N-n}Payoff_{VS}(0,T') - \frac{n}{N-n}Payoff_{VS}(0,T)$$

This equation establishes the additive property of variance swaps.

2 How to Make the Most of Variance Swaps?

The mechanism of variance swaps, as explained in our first section, is fairly easy: an investor pays the realized volatility over a given period of time, and another investor pays a variance strike, negotiated at the inception of the contract. One of the great advantages of variances swaps is that they are easy to replicate using only European options and futures. However, those remarks, important as they are, are purely descriptive: now we take a look at how variance swaps may be used as tool to leverage volatility.

2.1 Variance Swap Prices and the Variance Risk Premium

Looking at variance swap prices, i.e. the market strikes, is interesting as it allows us to identify a few statistical facts. Indeed, by definition, the variance strike of a variance swap is a measure of the anticipated volatility which will be realized over the lifetime of the swap, so it makes sense to compare those strikes with the usual implied volatilities computed thanks to European options prices.

Statistically, if we compare the strike price of six-month variance swaps with the six-month at-the-money implied volatility for European options on Eurostoxx 50, we see that the strike price is slightly superior to the implied volatility. If we denote $K_v^{Mkt}(6m)$ the market price, i.e. the variance strike, of six-month variance swaps, and $\sigma_i(100\%, 6m)$ the implied volatility of six-month, ATM European options, we have:

$$\frac{\sqrt{K_v^{Mkt}(6m)}}{\sigma_i(100\%, 6m)} \approx 1.1 > 1$$

The read must bear in mind that, since $K_v^{Mkt}(6m)$ is a variance strike, we have to take the square root of the value to be able to compare it with the implied volatility.

The same goes with realized volatility. We could have expected that variance swap strikes would have been a good estimation of the future realized volatility, however, when computing the regression of those strikes against the realized volatilities, we find that the slope is not equal to 1. Once again, the market strikes are above the volatility that is realized over the lifetime of the swap.

Those statistical observations can be restated this way: there exists a variance risk premium in the market. This is explained by the fact that being short a variance swap is riskier than being long the same instrument, insofar as the seller of a variance swap faces an potentially unlimited loss. The seller is then rewarded for taking this risk by being paid a variance strike which is higher than the volatility which is eventually realized at maturity.

This reality has been explored by Carr and Wu [3]. In a risk-neutral world, the variance risk premium, defined as the difference between the variance strike and the realized volatility, should be equal to 0. However the existence of risk aversion in the market explains why this difference is positive. We can underpin this assumption by studying the correlation between the skew and variance strikes.

The implied skew, measured as the spread between the 90 % and the 100 % implied volatilities, measures the risk aversion. Indeed, when investors are extremely risk-averse, they buy put options with strikes inferior to 100% in order to gain protection against sharp downturns. This explains why the implied volatility is higher for strikes below 100%: the higher the spread, the most important the risk aversion. Therefore it is interesting to compare the relationships between the skew and the variance risk premium. When performing the linear regression (Figure 1), we see that, the higher the skew, the higher the variance risk premium.

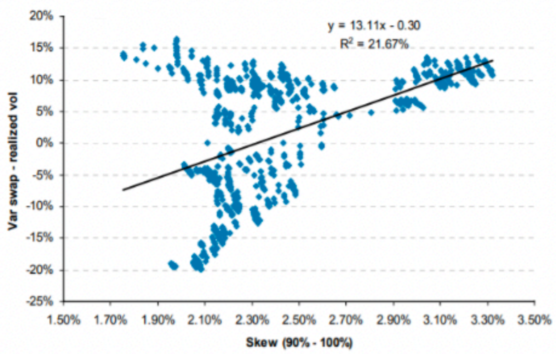

Figure 1: Regression between the one-year skew and the variance risk premium, measured with one-year variance swaps, for Eurostoxx 50

2.2 Volatility Strategies With Variance Swaps

2.2.1 The Arbitrage Between Implied and Realized Volatilities

One of the main aspects of the so-called volatility risk premium is that the implied volatility tends to be higher than the realized volatility, as shown by Figure 2. This fact is closely related to what we called in our previous section the "variance risk premium". First, the reader must not pay too much attention to the terminology: the expression "risk premia" was coined to refer to other-than-the-market sources of return, however, even though some of the risk premia are now well-known, their definition may be variable to subsume various aspects of markets' realities. This explains why, in our mind, we see the variance risk premium as a particular case of the more widespread "volatility risk premium".

Indeed, the fact that implied volatility is higher than the realized volatility can be explained with the same reasoning we set forth for the variance risk premium. The implied volatility measures the price of European option; it reflects the risk aversion that may be present among investors, thus leading to higher prices and higher implied volatilities.

Figure 2: Comparison between the implied volatility (blue) and the realized volatility (orange) for &P500

Variance swaps can be used to take advantage of this spread, or to take the opposite view if one expects the realized volatility to be higher than the implied volatility over a given period of time depending on the circumstances.

Being short a variance swap is a way of being paid the variance strike at maturity. Due to the definition of the strike, it can be seen as a measure of implied volatility, as it is an anticipation of the future realized volatility by market participants, like the usual implied volatility defined through European options prices. However variance strike prices, as we mentioned in subsection 2.1, are slightly superior to the usual ATM implied volatility, because you face a higher risk when shorting a variance swap, compared to investing in European options: a variance swap can lead an investor to an infinite loss. So variance swap is a powerful way of betting on the implied volatility, plus a risk premium based on the higher risk, against the realized volatility.

2.2.2 The Arbitrage of the Implied Volatility Term-Structure

For a given strike, the term-structure corresponds to the implied volatility with respect to the different maturities. This is typically an increasing curve: the higher the maturity, the higher the implied volatility. If an investor wants to take a view on the shape of the term structure, for instance if he or she expects the slope of the term structure to steepen in the next weeks or months, he or she can make such a bet thanks to either European options, or variance swaps.

If the investor decides to use options, he or she must invest in a calendar spread: if we consider two maturities $T_1 < T_2$, the investor buys a call with maturity T_2 and sells a call with maturity T_1. Indeed, if the slope steepens

$$\frac{\sigma_i(ATM, T_2) - \sigma_i(ATM, T_1)}{T_2 - T_1}$$

is going to increase: $\sigma_i(ATM, T_2)$ goes up whereas $\sigma_i(ATM, T_1)$ goes down. Since the vega of a call is positive, it means that the call with the longest maturity (T_2) will gain value, whereas the call with the shortest maturity (T_1) will lose value, thus increasing the overall position taken by the investor.

Another solution consists in relying on variance swaps. Once again, the investor expects a rise in the volatility term structure: he or she can take this view by being long a variance swap with a long maturity, and being short a variance swap with a short maturity.

As an example, let us assume that the long maturity is one year, and the short one is six months. The variance strike for the long maturity is $(25\%)^2$, but we expect the six-month variance swap to be priced at $(30\%)^2$ in six months. We also expect that the realized volatility σ will stay constant for the next year. If our bet is right, meaning that the term structure steepens, we wind up with the following payoff:

$$\left(\underbrace{(\sigma^2 - (25\%)^2)}_{long\ maturity} - \underbrace{(\sigma^2 - (35\%)^2)}_{short\ maturity} \right) \mathcal{N} > 0$$

2.2.3 Pairs Trading and Dispersion Trading

Pairs trading is a strategy involving two different stocks from the same sector: their volatility usually depends on the same factors, and so the spread is usually mean-reverting.

Let us assume we consider two similar stocks, for instance two European car-makers, such that the historical spread of six-month ATM implied volatilities is close to zero. If a spread appears, for instance if the implied volatility of the first car-maker spikes due to company-related circumstances, there is a trading opportunity which can be exploited thanks to variance swaps.

One way of taking advantage of such a situation is to short a variance swap on the stock whose implied volatility has just spiked, and to long a variance swap on the other stock. Being short the first variance swap, the investor receives at maturity a variance strike K_v^1 which is theoretically high, statistically higher than the implied volatility as seen in our previous section. However he or she has to pay the realized volatility: however, since the implied volatility is expected to decrease (mean-reversion of the spread), and since the implied volatility is statistically above the realized volatility, the investor expects the realized volatility to be contained. On the contrary, with his or her second variance swap, the investor pays a variance strike K_v^2, which may be assumed to be small since the implied volatility of the second stock has not spiked, and he or he receives the realized volatility, once again assumed to be close to its historical level.

Dispersion trading is an extension of pairs trading: it consists in buying the volatility of an index, while selling the volatility of its N constituents. Mathematically, dispersion is defined as:

$$D_t = \sum_{i=1}^{N} w_i^t \sigma_{i,t}^2 - \sigma_{Index,t}^2$$

where w_i^t is the weight of the i-th constituent of the index at date t, $\sigma_{i,t}$ its volatility at t, and $\sigma_{Index,t}$ the volatility at t of the overall index.

Dispersion can be reinterpreted in terms of correlations, denoted ρ in the following equations. Indeed:

$$D_t = \sum_{i=1}^{N} w_i^t \sigma_{i,t}^2 - cov_{Index,Index}^t$$

$$= \sum_{i=1}^{N} w_i^t \sigma_{i,t}^2 - \sum_{k=1}^{N} \sum_{j=1}^{N} w_k^t w_j^t cov_{k,j}^t$$

$$= \sum_{i=1}^{N} w_i^t \sigma_{i,t}^2 - \sum_{k=1}^{N} \sum_{j=1}^{N} w_k^t w_j^t \sigma_{k,t} \sigma_{j,t} \rho_{k,j}^t$$

$$= -2 \sum_{k<j} w_k^t w_j^t \sigma_{k,t} \sigma_{j,t} \rho_{k,j}^t$$

So the correlations between the various constituents play a significant role in the dispersion: as correlations tend to revert back towards their long-term mean, the same goes with dispersion. So it is possible to trade the volatility of an index against the volatilities of its constituents in a very similar way to pairs trading. Indeed, if the dispersion decreases, it can be interpreted as:

$$D_t = \underbrace{\sum_{i=1}^{N} w_i^t \sigma_{i,t}^2}_{down} - \underbrace{\sigma_{Index,t}^2}_{up}$$

Variance swaps provide an effective way of taking advantage of such a situation: the investor must go long a variance swap on the index, and short variance swaps on all the constituents of the index.

Conclusion

In this second paper about volatility investing and trading, we have provided the reader with an extensive presentation of variance swaps. Such contracts have indeed been developed in order to enable investors to gain pure exposure to volatility. Their success since more than twenty years is notably explained by two reasons. First they can be replicated through options and futures. Second they exhibit an additive property which makes them extremely practical when it comes to forward trading.

Variance swaps are also a powerful tool to harvest the volatility risk premium.

However, such contracts may also appear a bit restrictive. Indeed they only allow investors to take symmetric views on volatility: volatility is high or low, but the direction, upward or downward, of the underlying price moves is not taken into account. This is why other volatility products – the "third generation" – have been devised. This will be the focus of our third paper.

References

[1] P. Carr and D. Madan. Towards a theory of volatility trading. *Option Pricing, Interest Rates and Risk Management, Handbooks in Mathematical Finance*, 2001.

[2] M. Bouzoubaa and A. Osseiran. *Exotic Options and Hybrids: A Guide to Structuring, Pricing and Trading.* Wiley, 2010.

[3] P. Carr and L. Wu. Variance risk premia. *Courant Institute Working Paper*, 2005.

Nous sommes Canopee, un cabinet de conseil indépendant, spécialiste en Finance, DATA et transformation digitale.

Nous sommes l'*empowering ecosystem* ! Un écosystème en perpétuel mouvement.
Un écosystème qui donne le pouvoir à nos collaborateurs, nos projets et nos clients de se nourrir de l'émulation collective pour lui donner de la force.

Depuis 2009, nous intervenons dans les secteurs de la BFI de l'Asset Management, la banque privée ou de détail, les services financiers et l'assurance et cela auprès de clients tels que SG, HSBC,BNP Paribas ou encore ALLIANZ GI et AXA IM. Des écosystèmes réglementaires et spécifiques qui ont été nos premiers terrains de jeu.
C'est ici que nous avons su développer notre agilité, nos compétences, notre sens de l'engagement. Aujourd'hui, nous grandissons et continuons à nous investir et nous engager sur des écosystèmes spécifiques, techniques, mais diversifiés tels que l'industrie, le retail ou encore la pharma.
Toujours autour de nos 3 expertises que sont la finance, la data et la transformation digitale.

canopee
empowering *ecosystem*

www.canopee-group.com