WILMOTT TECHNICAL ARTICLE

PUBLISHED ON WILMOTT.COM IN JANUARY 2002



# GARCH AND VOLATILITY SWAPS

Alireza Javaheri, Paul Wilmott and Espen G. Haug

reza

Paul

Collector

Author details:

**Alireza Javaheri**,*
**Paul Wilmott**
and
**Espen G. Haug****

*Royal Bank of Canada, One Liberty Plaza, New York, NY 10006
Phone (212) 858-7353 or (718) 391-7283
Alireza.Javaheri@rbccm.com
**Paloma Partners
2 American Lane
Greenwich, CT 06836
Phone (203) 861 4838
EHaug@paloma.com

**Abstract**

This article discusses the valuation and hedging of Volatility Swaps within the frame of a GARCH(1,1) stochastic volatility model. First we use a general and exible PDE approach to determine the first two moments of the realized variance in a continuous or discrete context. Then we use this information to approximate the expected realized volatility via a convexity adjustment. Following this, we provide a numerical example using S&P500 data. Finally we describe a non-risk-neutral approach relying on the Central Limit Theorem for dealing with these volatility swaps practically.

Historical Note:
This paper is the result of a collaboration between members of the wilmott.com Forum, and initiated by the following thread "Garch and Vol Swaps" by reza: *"Hi, I am looking for methods of pricing Volatility Derivatives (e.g. Vol Swaps) using the GARCH model. So far I have seen one article (Heston & Nandi) but it's rather limited."*

# GARCH and Volatility Swaps

Alireza Javaheri,* Paul Wilmott† and Espen G. Haug

Contact person:

Alireza Javaheri
Royal Bank of Canada
One Liberty Plaza
New York, NY 10006
Phone (212) 858-7353 or
(718) 391-7283

or

Espen G. Haug
Paloma Partners
2 American Lane
Greenwich, CT 06836
Phone (203) 861 4838

January 11, 2002

*Alireza Javaheri, CFA works for Royal Bank of Canada.
†Paul Wilmott is a partner in Investment Analytics Inc.

# GARCH and Volatility Swaps[1]

**Abstract**

This article discusses the valuation and hedging of Volatility Swaps within the frame of a GARCH(1,1) stochastic volatility model. First we use a general and flexible PDE approach to determine the first two moments of the realized variance in a continuous or discrete context. Then we use this information to approximate the expected realized volatility via a convexity adjustment. Following this, we provide a numerical example using S&P500 data. Finally we describe a non-risk-neutral approach relying on the Central Limit Theorem for dealing with these volatility swaps practically.

[1]The authors would like to thank Omar Foda (University of Melbourne), Lennart Widlund (G&W Asset Management, Sweden), Jørgen Haug (Norwegian School of Economics), Robert Trevor (Macquarie University), Jin-chuan Duan (University of Toronto), Xingmin Lu (Royal Bank of Canada) and Christophe Bahadoran (Universite' Blaise-Pascal) for their very helpful comments and feedback. Robert Trevor also helped the authors with the numerical examples and the calibration process. The authors remain responsible for any errors in this paper. The discussions were mostly carried on www.wilmott.com/forum/. This is likely to be the first quantitative finance paper ever created over an internet forum.

# 1 Introduction

Variance and volatility swaps are useful products for volatility hedging and speculation. An interesting explanation for the takeoff for this market is given in Jim Gatheral's excellent course notes on variance and volatility swaps, NYU 2000 :

> "Variance swaps took off as a product in the aftermath of the LTCM meltdown in late 1998 when implied stock index volatility levels rose to unprecedented levels. Hedge funds took advantage of this by paying variance in swaps (selling the realized volatility at high implied levels). The key to their willingness to pay on a variance swap rather than sell options was that a variance swap is a pure play on realized volatility (no labor-intensive delta hedging or other path dependency is involved). Dealers were happy to buy vega at these high levels because they were structurally short vega (in the aggregate) through sales of guaranteed equity-linked investments to retail investors and were getting badly hurt by high implied volatility levels."

The market for variance and volatility swaps has since then been growing, and many investment banks and other financial institutions are now actively quoting volatility swaps on various assets: stock indexes, currencies, as well as commodities.

As a "warm up" to the valuation side we can suggest several papers: Carr and Madan (1998), Demeterifi, Derman, Kamal, and Zou (1999), Heston and Nandi (2000) and Brockhaus and Long (2000). A variance swap gives a payoff at maturity, which is equal to the difference between the realized variance over the swap period and the contract variance, multiplied by the notional. Similarly a volatility swap pays out the difference between the realized volatility covered by the swap and the contract volatility. In contrast to a variance swap a volatility swap is linear in pay-out for the realized volatility. As we soon will see a convexity adjustment is needed to value a volatility swap versus a variance swap. Market participants are used to trade volatility (through options), but not variance. In that respect, volatility swaps make more sense for most people. However, since variance swaps are easier to price, it seems like quants and institutions have preferred their clients to trade those (until now).

A stochastic volatility (or GARCH) model is needed for both the variance swap and the volatility swap. However the popular stochastic volatility models give us $\sigma^2$ (the volatility squared) and allow us to estimate the expectation $E[\sigma^2] = E[v]$, which is all that is needed for a variance swap. Since the price of any swap should be zero at issuance, the variance swap "delivery price" is $E[v] = E[\sigma^2]$. By the same token for a volatility swap the expected volatility is $E[\sigma] = E[\sqrt{v}]$. This is not available directly from the stochastic process for the common models. Still, one can use the Brockhaus and Long (2000) approximation (which is a Taylor expansion of order 2, of the square-root function on variable $v$ around the point $v_0 = E[v]$) to calculate

$$E[\sqrt{v}] \approx \sqrt{E[v]} - \frac{\text{Var}[v]}{8E[v]^{3/2}}.$$

See Appendix A for a quick proof of this. For a volatility swap we therefore need not only $E[v]$ but also $\text{Var}[v]$. We can also say that the term $\frac{\text{Var}[v]}{8E[v]^{3/2}}$ corresponds to the convexity adjustment or Jensen's Inequality correction for the square root function.

Moreover, it is actually the variance over the whole period covered by the swap which is important, and not only the variance at maturity. Let $I = \int_0^t v_u \, du$. We need to know $E[I]$ and $\text{Var}[I]$ to compute the convexity adjustment. Note that the convexity adjustment is only important when we are trying to find closed-form solutions for specific stochastic volatility processes. But even without talking about a specific stochastic volatility process in the Demeterifi, Derman, Kamal, and Zou (1999) paper it is the variance process that is replicated through a portfolio of plain vanilla calls and puts, and therefore there is still a need to go from a variance swap to a volatility swap using a convexity adjustment.

Note that we are using a continuous time process to calculate the convexity adjustment above. Still, we can use the discrete time process GARCH(1,1) to calculate $E[I_n]$, where $I_n$ is the discrete variance over the period. Even if the variance swap process is entirely discrete, the volatility swap process is not, unless we estimate $\text{Var}[I_n]$ using the continuous diffusion limit process and apply a finite difference scheme to the resulting PDE.

Later we will also discuss risk-neutrality and hedging of volatility swaps. We will suggest to drop risk neutrality and price via an expectation, a method that leads to a result that is more realistic and robust.

## 2 The Volatility Model

We will set up the problem for an arbitrary stochastic volatility model and then concentrate on one popular example: GARCH(1,1).

Supposing that

$$dv = f(v)dt + g(v)dX,$$

where $v$ is the 'variance' or, more precisely, the square of volatility, and $dX$ is a Wiener process, then we are interested in finding the following

**Problem A:**

1. $E[v]$
2. $\text{Var}[v]$

**Problem B:**

1. $E\left[\int_0^t v_u \, du\right]$
2. $\mathrm{Var}\left[\int_0^t v_u \, du\right]$

Here $E[\cdot]$ and $\mathrm{Var}[\cdot]$ are expectations at time $T$.

## 2.1 Problem A: Expectation and variance of $v$

The first two of these can be found by solving the Feynman-Kac backward equation

$$\frac{\partial}{\partial t} + \frac{1}{2} g(v)^2 \frac{\partial^2}{\partial v^2} + f(v) \frac{\partial}{\partial v} = 0 \tag{1}$$

subject to the relevant final conditions.

Using $F(v,t)$ to denote the expectation we are seeking and $G(v,t)$ the expected value of $v^2$ then both $F$ and $G$ satisfy Equation (1) with

$$F(v,T) = v$$

and

$$G(v,T) = v^2.$$

The variance of $v$ is then given by

$$G - F^2.$$

## 2.2 Problem B: Expectation and variance of $\int_0^t v_u \, du$

To address the second problem, we need to introduce a new state variable:

$$I_t = \int_0^t v_u \, du.$$

where $I$ can be seen as the realized variance. More precisely $I$ is the variance over the life of the contract, as opposed to $v$ the instantaneous variance at a point in time. To be technically correct, we should say

$$I_t = \frac{1}{T} \int_0^t v_u \, du.$$

This way on the issue date of the swap, the realized volatility will coincide with the instantaneous volatility and we can mark-to-market the swap during its life, which would allow us to have American swaps as well. However this division by $T$ does not change our equations and we will not include it for simplicity reasons. Now we must solve

$$\frac{\partial}{\partial t} + \frac{1}{2} g(v)^2 \frac{\partial^2}{\partial v^2} + f(v) \frac{\partial}{\partial v} + v \frac{\partial}{\partial I} = 0 \tag{2}$$

subject to the relevant final conditions.

Using $F(v, I, t)$ to denote the expectation of $I$ and $G$ the expected value of $I^2$ then both $F$ and $G$ satisfy Equation (2) with

$$F(v, I, T) = I$$

and

$$G(v, I, T) = I^2.$$

The variance of $I$ is then given by

$$G - F^2.$$

## 3 Special case: GARCH(1,1)

The model for the variance in a continuous version is specifically

$$dv = \kappa(\theta - v)dt + \gamma v \, dX.$$

where $\kappa$ is the speed of mean reversion, $\theta$ is the mean reversion level, and $\gamma$ is the volatility of volatility, more precisely, the volatility of the square of volatility. The discrete version of the GARCH(1,1) process is described in Engle and Mezrich (1995)

$$v_{n+1} = (1 - \alpha - \beta)V + \alpha u_n^2 + \beta v_n$$

where $V$ is the long-term variance, $u_n$ is the drift-adjusted stock return at time $n$, $\alpha$ is the weight assigned to $u_n^2$, and $\beta$ is the weight assigned to $v_n$.[2] Further we have the following relationship

$$\theta = \frac{V}{dt}$$

$$\kappa = \frac{1 - \alpha - \beta}{dt}$$

$$\gamma = \alpha\sqrt{\frac{\xi - 1}{dt}}$$

[2]The GARCH(1,1) model implies that the stock process and the volatility process contain two uncorrelated Brownian Motions. In an NGARCH process, described by Engle and Ng (1993), we have

$$v_{n+1} = (1 - \alpha - \beta)V + \alpha(u_n - c)^2 + \beta v_n$$

where $c$ is another parameter to be estimated and creates the correlation between the two processes. However this will not affect our discussion at all and therefore we will use the GARCH(1,1) model throughout this paper. The NGARCH process is discussed for instance in Ritchken and Trevor (1997).

where[3] $\xi$ is the Pearson kurtosis[4] (fourth moment) of $u[n]$.

A stochastic volatility model (or GARCH) gives us only $dv$ where $v_t$ (or $v_n$) is the variance and not the volatility. So, for both the continuous version above or the discrete version we do need the Taylor Expansion to find the value of a volatility swap:

$$E[\sqrt{v}] \approx \sqrt{E[v]} - \frac{\mathrm{Var}[v]}{8E[v]^{3/2}}.$$

## 3.1 Problem A

The problem for $F$ can be written as

$$\frac{\partial F}{\partial t} + \frac{1}{2}\gamma^2 v^2 \frac{\partial^2 F}{\partial v^2} + \kappa(\theta - v)\frac{\partial F}{\partial v} = 0$$

with $F(v, T) = v$. This has the simple solution

$$F(v,t) = \theta + e^{-\kappa(T-t)}(v - \theta).$$

The solution for $G(v,t)$ is also easily found to be

$$\begin{aligned} G(v,t) &= \frac{2\kappa^2\theta^2}{\gamma^2 - \kappa}\left(\frac{e^{(\gamma^2-2\kappa)(T-t)} - 1}{\gamma^2 - 2\kappa} + \frac{e^{-\kappa(T-t)} - 1}{\kappa}\right) \\ &\quad + \frac{2\kappa\theta}{\gamma^2 - \kappa}\left(e^{(\gamma^2-2\kappa)(T-t)} - e^{-\kappa(T-t)}\right) v + e^{(\gamma^2-2\kappa)(T-t)} v^2. \end{aligned}$$

From $F$ and $G$ we can calculate the variance of $v$ as indicated above.

## 3.2 Problem B

The problem for $F$ can be written as

$$\frac{\partial F}{\partial t} + \frac{1}{2}\gamma^2 v^2 \frac{\partial^2 F}{\partial v^2} + \kappa(\theta - v)\frac{\partial F}{\partial v} + v\frac{\partial F}{\partial I} = 0$$

with $F(v, I, T) = I$.

---

[3]The $\gamma$ in the Engle and Mezrich (1995) paper is $\gamma = \alpha\sqrt{(\xi - 1)dt}$ which is different from our defintion. This is likely due to a small typo in their paper. See also Nelson (1990).

[4]Pearson kurtosis is Fisher kurtosis plus three. The normal distribution has a Pearson kurtosis of 3 (Fisher kurtosis of 0), called mesokurtic. Distributions with Pearson kurtosis larger than 3 (Fisher higher than 0) are called leptokurtic, indicating higher peak and fatter tails than the normal distribution. Pearson kurtosis smaller than 3 (Fisher lower than 0) is termed playakurtic. Before calculating kurtosis from asset prices make sure you know if the software you are using returns Pearson or Fisher kurtosis.

The solution for $F(v, I, t)$ is now

$$F(v, I, t) = \theta \left( T - t + \frac{e^{-\kappa(T-t)} - 1}{\kappa} \right) + \frac{1}{\kappa} \left( 1 - e^{-\kappa(T-t)} \right) v + I.$$

Similarly, $G(v, I, t)$ has solution

$$G(v, I, t) = f(t) + g(t)v + h(t)v^2 + l(t)I + n(t)vI + I^2$$

with

$$\begin{aligned} f(t) \quad &= \quad \theta^2(T-t)^2 - \frac{4\theta^2(\gamma^2 - \kappa)}{\kappa(\gamma^2 - 2\kappa)} \left( T - t + \frac{e^{-\kappa(T-t)} - 1}{\kappa} \right) \\ &\quad - \frac{4\theta^2\kappa^2}{(\gamma^2 - \kappa)^2(\gamma^2 - 2\kappa)} \left( \frac{1 - e^{(\gamma^2 - 2\kappa)(T-t)}}{(\gamma^2 - 2\kappa)} + \frac{1 - e^{-\kappa(T-t)}}{\kappa} \right) \\ &\quad - \frac{2\theta^2(\gamma^2 + \kappa)}{\gamma^2 - \kappa} \left( e^{-\kappa(T-t)} \frac{T-t}{\kappa} + \frac{1}{\kappa^2}(e^{-\kappa(T-t)} - 1) \right) \end{aligned}$$

$$\begin{aligned} g(t) \quad &= \quad \frac{2\theta}{\kappa}(T-t) - \frac{4\theta(\gamma^2 - \kappa)}{\kappa^2(\gamma^2 - 2\kappa)} \left( 1 - e^{-\kappa(T-t)} \right) \\ &\quad + \frac{4\theta\kappa}{(\gamma^2 - \kappa)^2(\gamma^2 - 2\kappa)} \left( e^{(\gamma^2 - 2\kappa)(T-t)} - e^{-\kappa(T-t)} \right) + \frac{2\theta(\gamma^2 + \kappa)}{\kappa(\gamma^2 - \kappa)}(T-t)e^{-\kappa(T-t)} \end{aligned}$$

$$h(t) = \frac{2}{\kappa(\gamma^2 - 2\kappa)} \left( e^{(\gamma^2 - 2\kappa)(T-t)} - 1 \right) - \frac{2}{\kappa(\gamma^2 - \kappa)} \left( e^{(\gamma^2 - 2\kappa)(T-t)} - e^{-\kappa(T-t)} \right)$$

$$l(t) = 2\theta \left( T - t + \frac{e^{-\kappa(T-t)} - 1}{\kappa} \right)$$

$$n(t) = \frac{2}{\kappa} \left( 1 - e^{-\kappa(T-t)} \right)$$

## 3.3 Explicit Expression for $v_t$

It would be interesting to use an alternative method to calculate $F(v, t)$ and the other above quantities. We could calculate $v_t$ directly by solving the stochastic differential equation (SDE)

$$dv = \kappa(\theta - v)dt + \gamma v \, dX.$$

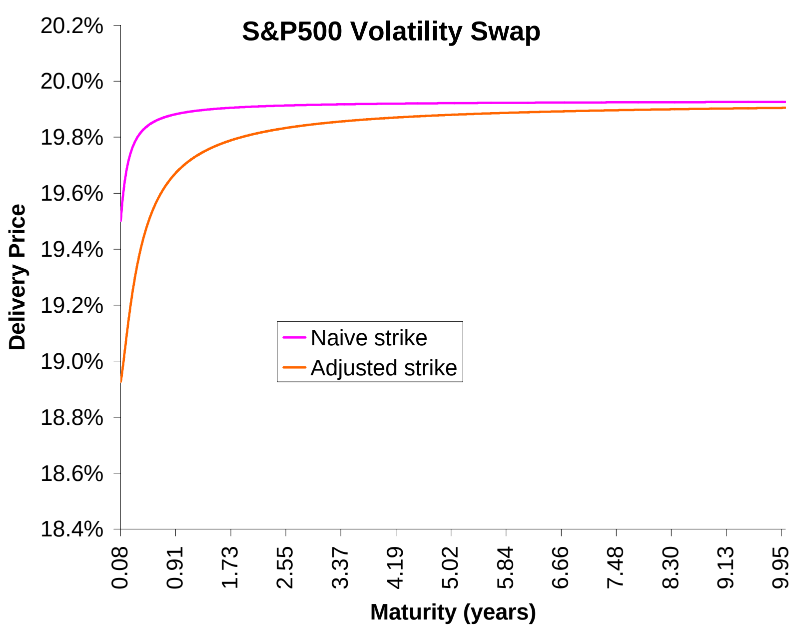
This equation can be solved via a variation of constants technique and the solution is

$$v_\tau = v_0 e^{-(\kappa+\frac{\gamma^2}{2})\tau+\gamma X_\tau} + \kappa\theta \int_0^\tau e^{-(\kappa+\frac{\gamma^2}{2})(\tau-u)+\gamma(X_\tau - X_u)}\, du$$

with $\tau = T - t$. We can see immediately that the resulting $F(v,t)$ is the same as the one in is section (3.1) using the well known equation $E(e^{\gamma X_\tau}) = e^{\frac{1}{2}\gamma^2\tau}$. A short summary of how to calibrate the GARCH(1,1) model is given in appendix B.

## 3.4 Numerical examples

Before we move on let's take a look at some numeric examples. We are calibrating the GARCH parameters from five years of daily historic S&P500 (SPX) prices (from 1996/10/01 to 2001/09/28). This gives us a long term variance of $V = 0.00015763$ (so the annualized long-term volatility is around 19.93%) and a kurtosis of 5.81175. The discrete GARCH(1,1) parameters are $\alpha = 0.127455$, $\beta = 0.7896510$. Based on business daily data and therefore $dt = 1/252$, the diffusion limit parameters are $\theta = 0.0397224$, $\kappa = 20.889356$ and $\gamma = 4.4382085$. We take for inputs $v = 0.0361$ (corresponding to a volatility of 19%) and $I = 0$ (on the issue-date of volatility swap). Figure 1) illustrates the strike of a volatility swap from a GARCH(1,1) model with and without convexity adjustment. Figure 2) shows the convexity adjustment itself, which is the difference between the naive strike and the convexity adjusted-strike.



# WWW.WILMOTT.COM BOOKSHOP

## *The Complete Guide to Option Pricing Formulas* by Espen Haug

**Jacket text**

When pricing options in today's fast-action markets, experience and intuition are no longer enough. To protect your carefully planned positions, you need precise facts and tested information that has been proven time and again. The Complete Guide to Option Pricing Formulas is the first and only authoritative reference to contain every option tool you need, all in one handy volume: Black-Scholes, two asset binomial trees, implied trinomial trees, Vasicek, exotics.

Many important option pricing formulas are accompanied by computer code to assist in their use, understanding, and implementation. This invaluable, one-of-a-kind reference work gives you:

- A complete listing of key option formulas, all delivered in an easy-to-use dictionary format
- Commentary that explains key points in the most important and useful formulas
- Valuable software and ready-to-use programming code that enhances your understanding of option pricing models and their practical implementations
- Practitioner-orientated formulas, and highlights of the latest option pricing research from major institutions worldwide
- Pricing advances on commodity options like the Miltersen and Schwartz Model, exotic options such as extreme spread options and implied trinomial trees, and much more!

Professionals who use options must have immediate access to reliable and complete option pricing formulas and information. The Complete Guide to Option Pricing Formulas, an invaluable guide for both experienced users and those learning how to use the tolls of valuation, is the first book to place all of the research and information you need at your fingertips.

**Contents**

1. Plain Vanilla Options

2. Exotic Options

3. Numerical Methods in Options Pricing

4. Interest Rate Options

5. Volatility and Correlation

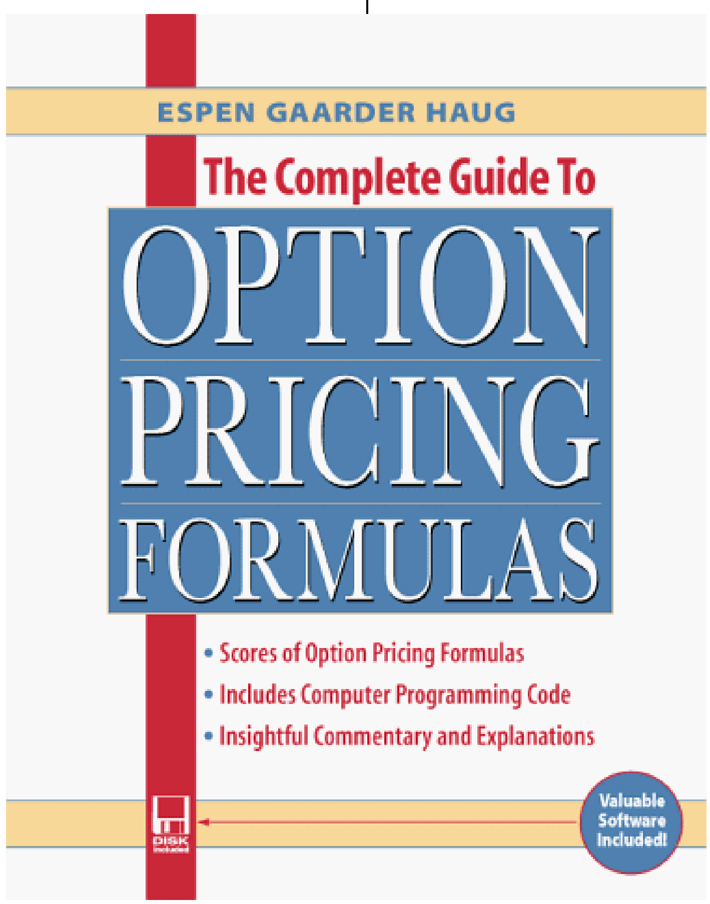
6. Some Useful Formulas

A. Distributions

B. Partial Derivatives of the Black-Scholes

C. The Option-Pricing Software

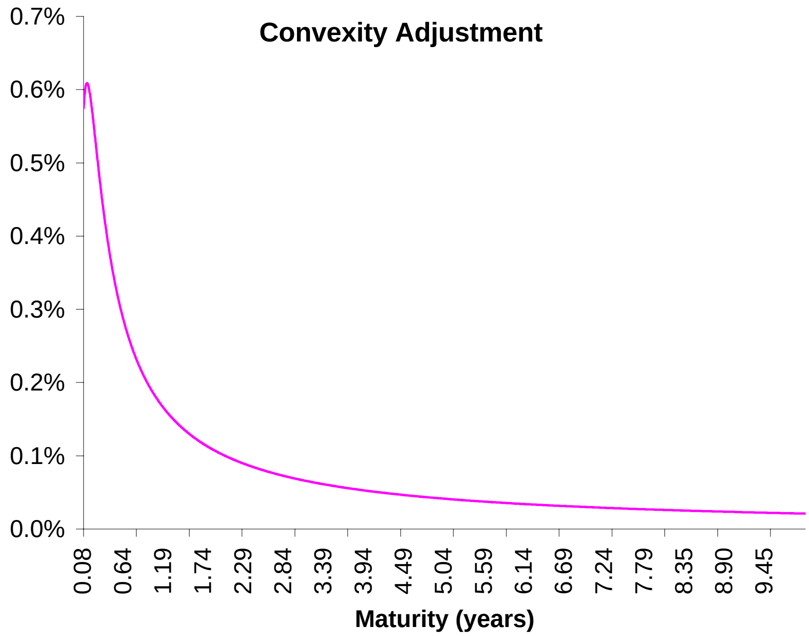
Bibliography




If you like this article from wilmott.com, you might consider purchasing this excellent book also written by Espen Haug.

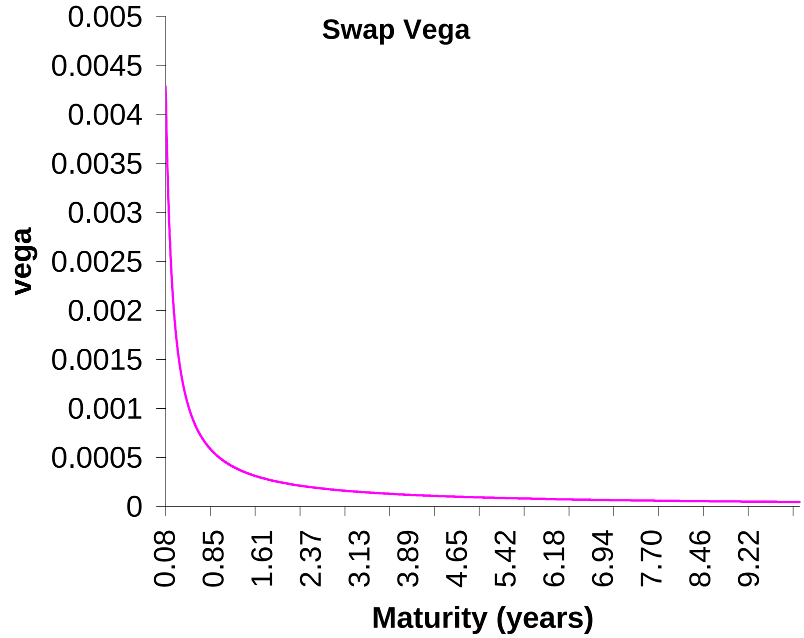
Published in 1998 by McGraw Hill

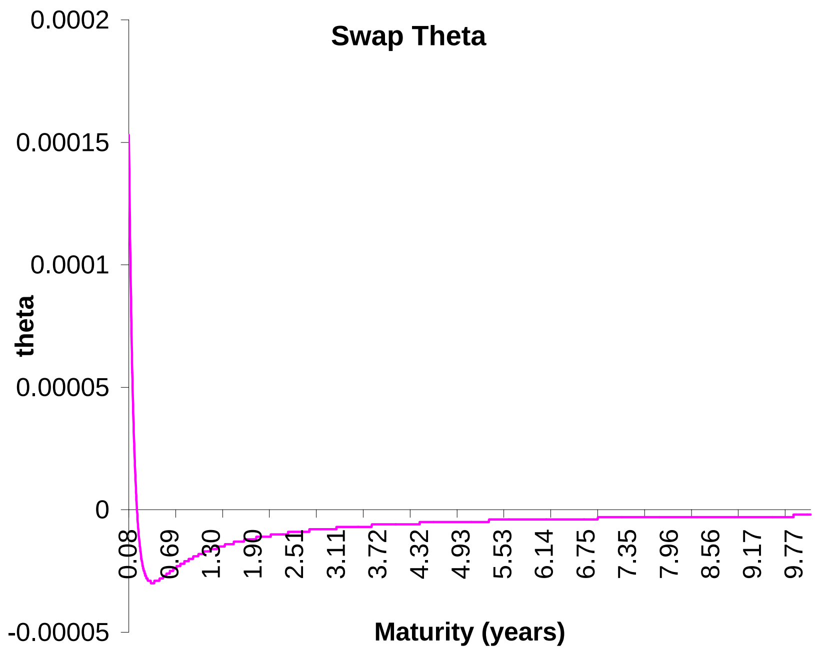
Visit the Quantitative Finance bookshop at **www.wilmott.com** to order




From Figure 2 we can easily see how the convexity-adjustment is decreasing with swap maturity. The intuition behind this is mainly that the volatility of the volatility over a long period of time is low.

For market participants there would typically be of great interest to know how sensitive a volatility swap is to change in volatility. Figure 3 shows the sensitivity of the strike/ fair value of a volatility swap with respect to one point change in the input volatility (the vega). Similarly Figure 4 illustrates the one day theta of a volatility swap (change in the strike of the volatility swap as time to maturity decreases by one day).


Swap Vega
vega
0.005
0.0045
0.004
0.0035
0.003
0.0025
0.002
0.0015
0.001
0.0005
0
0.08
0.85
1.61
2.37
3.13
3.89
4.65
5.42
6.18
6.94
7.70
8.46
9.22
Maturity (years)


Swap Theta
theta
0.0002
0.00015
0.0001
0.00005
0
-0.00005
0.08
0.69
1.30
1.90
2.51
3.11
3.72
4.32
4.93
5.53
6.14
6.75
7.35
7.96
8.56
9.17
9.77
Maturity (years)

# 4 Volatility swaps and risk neutrality

There are several problems with the traditional risk-neutral hedging concept for variance and volatility swaps. Demeterifi, Derman, Kamal, and Zou (1999) suggest to replicate a variance swap with a forward contract and an infinite sum of calls and puts, which they argue can be approximated with a finite number of options. This method has several pitfalls and is probably not practical to implement. Realistically, we can't hedge this product, so we have to appeal to the concept of market price of risk. Or do we? There are several reasons to avoid models that require the input of market price of risk, which is very difficult to estimate. Even at its best, the market price of risk is very unstable[5], see Wilmott (2000) chapter 36. Here we will suggest to drop risk neutrality and price via an expectation. So $E[\sqrt{v}]$ or $E[v]$ become our real expected payoff. We stress 'real'. But this payoff isn't guaranteed, it's risky. The payoff may be more or may be less. This is where the $\mathrm{Var}[\sqrt{v}]$ or $\mathrm{Var}[v]$ comes in. If we are now using the real (not the risk-neutral) process for volatility in a mean-variance framework (see Ahn, Arkell, Choe, Holstad, and Wilmott (1999), and Wilmott (2000) chapter 36) the delivery price $K$ of a volatility swap will be

$$K = E[\sqrt{v}] - \lambda \mathrm{Stdv}[\sqrt{v}]$$

where $\lambda$ measures our own personal risk aversion, and the minus sign would change to a plus sign if we are short volatility instead of being long. The same concept naturally holds for $E[I]$ and $E[\sqrt{I}]$.

Back to the $\lambda$ parameter, market participants will get different prices depending upon their risk-aversion parameter (much like utility theory). To give some intuition on this parameter:

- Choose $\lambda = 0$ and you never make money, nor lose it (on average, that is, over many deals).
- Choose $\lambda$ large and you will find that your price is outside the market's hence you never do any deals and again don't make any money.
- Choose $\lambda$ small and you will make a little bit of profit on average, and maybe do quite a few deals. So, at this stage, there is an optimization to be done. You can think of this optimization as making the most from the Central Limit Theorem.

Concerning the Central Limit Theorem and the mean-variance framework for a given $\mu = E[v]$ and $\sigma = \sqrt{\mathrm{Var}[v]}$ and our personal (fixed) degree of risk-aversion $\lambda$, if we are long volatility, then the delivery-price will be

$$K = \mu - \lambda\sigma$$

Now if we have $n$ transactions on with the same law and same $\mu$ and $\sigma$ and supposing

[5] Where is the market price of volatility risk in the Heston and Nandi (2000) model? In the drift of the volatility.

1. $n$ is large enough.
2. and we assume the $v_1, v_2, \ldots, v_n$ are independent.

We can now apply the Central Limit Theorem and we can say

$$V_n = \frac{v_1 + \ldots + v_n}{n} \approx N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$$

where $N()$ the normal distribution. So our decision to make a trade or not would depend on $K_n = \mu - \lambda \frac{\sigma}{\sqrt{n}}$ being within the market bid/ask interval for the delivery-price or not. As the number of deals $n$ goes higher $K_n$ approaches $K$ and gets a better chance to actually be within this bid/ask interval.

To summarize, the suggested approach has several advantages over other proposed solutions:

1. It doesn't depend on the market price of risk.
2. The price can vary from person to person. In a world where everyone agrees on the same price there's not much point in trading.
3. It is nonlinear, therefore economies of scale apply. This also gives us a bid-offer spread on volatility swaps and would allow optimal semi-static hedging (see Wilmott (2000) chapter 32).

# 5 Conclusion

We have given you two different methods to compute $E[v_t]$ and $\text{Var}[v_t]$ as well as $E[I]$ and $\text{Var}[I]$. From these we can not only value a variance swap but also determine the convexity adjustment for a volatility swap. The first method we used is the PDE method described in section (2.1). The main strength of this method is its flexibility. With the PDE method we can basically find the expected variance and variance of the variance for any given function[6], either explicitly, or for more complex cases by using finite difference methods.

In the second method we solved the stochastic differential equation using the usual (non-stochastic) differential-equation tricks in combination with Ito's lemma to come up with closed form solutions. Closed form solutions offer additional insight and intuition that are not readily available from PDE methods.

A drawback of most stochastic volatility solutions, whether PDE or closed form based, is that they typically assume continuous monitoring of the volatility. This problem is also discussed by Carr and Corso (2001). In practice all variance and volatility swaps are based on discrete monitoring, typically some type of daily close (fixing) price. Even if we assume constant volatility, for instance geometric Brownian motion, the volatility of volatility could be high simply because of discrete monitoring.[7] In this paper we have the solution for the

[6]For instance $E[e^{-av_t}]$ poses no problems.

[7]See for example Haug (1997) pages 169-170 and Wilmott (2000) pages 299-301.

expected continuous time variance $E[I]$ and the expected variance in a discrete GARCH(1,1) model

$$E[I_n] = V + \frac{1-(\alpha+\beta)^n}{(1+n)[1-(\alpha+\beta)]}(v_0 - V)$$

where $I_n$ is the discrete variance covering the period[8]. However for the variance of the variance we only have the solution for the continuous time GARCH process. In order to obtain $\mathrm{Var}[I_n]$ we could apply a finite difference scheme to the PDE method of section (2.1) and therefore obtain an entirely discretely monitored volatility swap pricing model.

# References

AHN, H., R. ARKELL, K. CHOE, E. HOLSTAD, AND P. WILMOTT (1999): "Optimal Static Vega Hedge," *MFG working paper, Oxford University.*

BROCKHAUS, O., M. FARKAS, A. FERRARIS, D. LONG, AND M. OVERHAUS (2000): *Equity Derivatives and Market Risk Models.* Risk Books.

BROCKHAUS, O., AND D. LONG (2000): "Volatility Swaps made simple," *Risk Magazine*, January.

CARR, P., AND A. CORSO (2001): "Commodity Covariance Contracting," *Energy and Power Risk Management*, April, http://www.math.nyu.edu/research/carrp/papers/.

CARR, P., AND D. MADAN (1998): *Towards a Theory of Volatility Trading.* In the book: Volatility, Risk book publications, http://www.math.nyu.edu/research/carrp/papers/.

DEMETERIFI, K., E. DERMAN, M. KAMAL, AND J. ZOU (1999): "More Than You Ever Wanted to Know About Volatility Swaps," *Working Paper, Goldman Sachs, http://www.gs.com/qs/doc/volswaps.pdf.*

ENGLE, R. F., AND J. MEZRICH (1995): "Grappling with GARCH," *Risk Magazine*, September.

ENGLE, R. F., AND V. K. NG (1993): "Measuring and Testing the Impact of News on Volatility," *Journal of Finance*, 48(5), 1749–1779.

GATHERAL, J. (2000): "Variance Swaps," *Case Studies in Financial Modeling Course Notes, NYU Courant Mathematical Institute*, Spring.

HAUG, E. G. (1997): *The Complete Guide To Option Pricing Formulas.* McGraw-Hill, New York.

[8] We also have $E[v_n] = V + (\alpha+\beta)^n (v_0 - V)$.

HESTON, S. L., AND S. NANDI (2000): "Derivatives on Volatility: Some Simple Solutions Based on Observables," *Working Paper 2002-20 Federal Reserve Bank of Atlanta, http://www.frbatlanta.org/publica/work_papers/wp00/wp0020.pdf*.

HULL, J. (2000): *Option, Futures, and Other Derivatives*. Prentice Hall.

NELSON, D. B. (1990): "ARCH Models as diffusion Approximations," *Journal of Econometrics*, 45, 7–38.

PRESS, W. H., S. A. TEUKOLSKY, W. T. VETTERLING, AND B. P. FLANNERY (1993): *Numerical Recipes in C*. Chapter 15, Cambridge University Press, http://www.nr.com.

RITCHKEN, P., AND R. TREVOR (1997): ""Pricing Options under Generalized GARCH and Stochastic Volatility Processes," *CMBF Papers*, 19, http://www.mafc.mq.edu.au/MAFCpapers/papers.htm.

WILMOTT, P. (2000): *Paul Wilmott on Quantitative Finance*. Wiley.

# Appendix A: Quick proof of the Brockhaus-Long approximation

Here is a quick proof of the Brockhaus and Long (2000) approximation, see also Brockhaus, Farkas, Ferraris, Long, and Overhaus (2000). Given $F$ the square-root function

$$F(x) = x^{1/2}$$

We have

$$F'(x) = \frac{1}{2x^{1/2}}$$

$$F''(x) = -\frac{1}{4x^{3/2}}$$

The second order Taylor expansion for $F(x)$ around $x_0$ gives

$$\begin{aligned} F(x) &\approx F(x_0) + F'(x_0)(x - x_0) + \frac{1}{2}F''(x_0)(x - x_0)^2 \\ &\approx x_0^{1/2} + \frac{x - x_0}{2x_0^{1/2}} - \frac{1}{8}\frac{(x - x_0)^2}{x_0^{3/2}} \\ &\approx \frac{x + x_0}{2x_0^{1/2}} - \frac{(x - x_0)^2}{8x_0^{3/2}} \end{aligned}$$

Now if we take $x = v$ and $x_0 = E[v]$ we will have

$$\sqrt{v} \approx \frac{(v + E[v])}{2\sqrt{E[v]}} - \frac{(v - E[v])^2}{8E[v]^{3/2}}$$

and taking Expectations on both sides

$$E[\sqrt{v}] \approx \frac{E[v] + E[v]}{2\sqrt{E[v]}} - \frac{E[(v - E[v])^2]}{8E[v]^{3/2}}$$

or

$$E[\sqrt{v}] \approx \sqrt{E[v]} - \frac{\mathrm{Var}[v]}{8E[v]^{3/2}}$$

which is the Brockhaus-Long approximation. So all we need to find the convexity adjustment for a volatility swap is $E[v]$ and $\mathrm{Var}[v]$ for the given choice of process. In our case GARCH(1,1).

# Appendix B: Calibration to historical data

For the calibration of the GARCH(1,1) model we can use a maximization of likelihood ratio for the stock return density and we will get the optimum $\alpha$ and $\beta$, see Engle and Mezrich (1995) and Hull (2000). Then we will deduce $\theta$, $\kappa$, $\gamma$ from the former parameters together with our choice of $dt$. One would try to minimize

$$\sum_{n=1}^{N}\left(\ln(v_n)+\frac{u_n^2}{2v_n}\right)$$

with

$$v_n=(1-\alpha-\beta)V+\alpha u_{n-1}^2+\beta v_{n-1}$$

where $u_n$ is the drift-adjusted stock return at time $n$ and $N$ the total number of observations. One way would be to use the Fletcher-Reeves-Pollak-Ribiere minimization method, see Press, Teukolsky, Vetterling, and Flannery (1993). Note that this method is non-constrained and would "blow-up" if used like this, so inside the function to be minimized (the above sum) one would need to include the constraints $0<\alpha<1$ , $0<\beta<1$, $0<\alpha<1-\beta$.

Note that, given our stochastic process $dv=\kappa(\theta-v)dt+\gamma v dX$, if $\gamma$ the volatility of the volatility is too large, i.e. if the term $\kappa-\frac{1}{2}\gamma^2$ becomes negative, then our variance term and therefore our convexity adjustment will diverge for $T-t$ large enough. In order to make sure that $\kappa-\frac{1}{2}\gamma^2>0$ we can add to our usual GARCH(1,1) optimization constraints, the condition $1-\alpha-\beta-\frac{\xi-1}{2}\alpha^2>0$.

# Appendix C: Calibration to market data

If we want to calibrate the model against current market prices (as opposed to historic prices) we need to make the parameter $\theta$ representing the long-term variance, time dependant. Therefore instead of estimating $V$ from the historic prices and deduce $\theta$ from there only once, we recalculate them dynamically. The rest of what we did however remains the same.