Risk Application for VIX Derivatives Using a Multi Factorial Model Within the No-Arbitrage Framework

Marc-André Picard
Université Laval
Essai de Maîtrise 12 crédits

17 mars 2015
Special Thanks

Before jumping in the heart of the matter, there are people in merit of an honourable mention without which this work would have been impossible. My family, friends and professors have had a major role in keeping me on track and gave me the ambition and the will to see this project through the end.

I am immensely grateful to have had the chance of being a student of Professor Van Son Lai who not only brought me intuition and insights but also had me develop a passion for finance.

I would also like to thank M. Philippe Bélanger for accepting to review my work.

I am also grateful to my mother and father, Lise and Gilles, as well as my beloved Chelsea for their unconditional support and to the friends that have been by my side during this time.

Thank You.
Contents

1 Introduction 4

2 The VIX Index 6
   2.1 The Origins ................................................. 6
   2.2 The History .............................................. 9
   2.3 The Methodology ......................................... 11
   2.4 The Purpose .............................................. 13

3 VIX ETPs 20
   3.1 Specific Case: Short-Term Volatility Tracking Index ETN (VXX) 22

4 Volatility Derivatives 26
   4.1 Black-Type models ....................................... 27
       4.1.1 Whaley (1993) .................................. 27
       4.1.2 Carr & Lee (2007) ............................... 27
   4.2 Equilibrium Models ..................................... 29
       4.2.1 Grünbichler & Longstaff (1996) ............... 29
       4.2.2 Zhang & Zhu (2006) ............................ 33
       4.2.3 Sepp (2008) ...................................... 35
       4.2.4 Lin & Chang (2009) .................... 36
   4.3 No Arbitrage Models .................................... 37
       4.3.1 Dupire (1993) .................................. 37
       4.3.2 Buhler (2006) .................................. 40
       4.3.3 Huskaj & Nossman (2013) ................... 41
       4.3.4 Cont & Kokholm (2013) .................... 43

5 Model Specifications 45
   5.1 Multi-Factorial No-Arbitrage Model for VIX Futures .... 47
   5.2 Curve Dynamics ........................................ 48
   5.3 Forward Variance Volatility ............................ 49
   5.4 VIX futures VaR ........................................ 50

6 Results 51
   6.1 Model Validation ...................................... 51
       6.1.1 Input vs. Output ................................ 52
       6.1.2 European VIX Option Valuation ............ 60
   6.2 Distributions and Value-at-Risk ........................ 63

7 Discussion and further improvements 69

8 Conclusion 71
1 Introduction

Developed by Robert E. Whaley in 1993, the VIX volatility index has become a trademark product of the Chicago Board Option Exchange (CBOE) and is currently used throughout today’s market. The VIX Index represents the market’s expectation of the S&P 500 Index annualized volatility over a 30 days period.

The investor’s interest for the VIX has been growing since its creation. To give a slight idea, the total open interest on January 3rd 2012 for the January, February, and March futures contracts was 88297. This represents a market value close to 230 million USD. The total traded volume for the three contracts on that day was 23808 which represent approximately 62 million USD.\footnote{Data from CBOE.com.}

It is important from the get go to understand that the VIX is not a traded asset and is a forward looking measure in a sense that it represents the anticipation of the market volatility in the next 30 days or in other terms the cost of portfolio insurance with a 30 days maturity. Many developments and improvements have been done since its inception to refine the calculation methods for the VIX but also a great deal of research to address the problem of VIX derivative pricing.

Since the CBOE released the VIX futures in 2004 and the VIX options in 2006, the VIX derivatives have been growing in popularity not only on the Exchange but also in the financial literature. While these products provide a simple and cost efficient way to hedge against volatility risk, VIX derivatives have a number of properties which make the pricing become problematic compared to equity index derivatives and even more so when looking at the multiple products that have become available in the recent years.

Since January 2009, smaller investors have been able to benefit from those volatility hedging tools through a wide variety of Exchange Traded Products (ETPs) that track the VIX Futures return index along with their respective set of derivatives. These volatility ETPs constructed from VIX futures of different maturities have grown in popularity since their inception and since the market has seen the appearance of a wide array of these products such as short, mid and long term volatility tracker with their inverse as well as double and triple capped trackers.

The VIX futures and options present a challenge in themselves when it comes to valuation. The main goal of this research will however not be on the pricing methods but will instead attempt to model and quantify the risk of holding VIX products in a portfolio. Recent studies show an interesting phenomenon in which the VIX futures are bound to converge towards the spot which makes their returns dependant on the shape of the actual term-structure.\footnote{Referred to as Roll-Down Effect.} Therefore modeling the dynamics of the VIX futures term-structure itself by taking the prices of the VIX futures available on the market for granted should allow to
capture more realistic risk measures and distributions than classic risk management methods.

In order to judiciously choose a suiting model for our application, we will start by studying the VIX index in itself as well as VIX derivatives and exchange traded products. We will then take a look at the strengths and weaknesses of volatility and variance models in the past literature which will lead us to our model choice of a multi-factorial no-arbitrage forward variance model that relies on readily available market information for its volatility specifications and uses principal component analysis method (PCA) to model the curve’s dynamics.

By modelling the forward variance curve, we are then able to obtain a process for the VIX futures that is free of arbitrage and that fits the initial VIX futures curve’s shape perfectly. By using quoted VIX options and defining a correlation structure between the forward variance maturities, we are able to back-out the volatility specifications for the model. We are then able quantify the risk of holding a VIX futures position in a portfolio while taking into account the roll-down effect associated with the shape of the VIX futures curve while allowing the curve to move more freely by using multiple factors to drive its dynamics.

Our results show that when comparing our multi-factorial model to a one factor model, there is significant difference in the distribution of the roll-down effect of the futures. The differences in the roll-down distribution when using a multi-factorial model has a slight impact on the skewness and kurtosis of the distribution when holding a position in a single VIX futures contract but becomes far more apparent when holding a portfolio of VIX futures of different maturities. This in turn impacts significantly the Value-at-Risk (VaR) of a portfolio holding multiple maturities.

Our results also show that the use of PCA to model the curve’s dynamics reduces the number of factor needed to drive the curve while still capturing most of the possible curve scenarios (93% vs. 71% when using only one factor). When comparing to more classical methods such as correlated random samples generation using Cholesky factorisation that would allow to capture 100% of the possible curve’s movements but where we would need one factor per simulated maturity, the PCA reduces significantly the computing time (approximately 38%) when simulating the entire VIX curve that usually includes the spot VIX along with six additional forward maturities.

Section 2 presents a study of the VIX Index including its origins and history as well as explaining its purpose and methodology. Section 3 takes a look at the different VIX ETPs and their properties. Section 4 presents VIX derivatives along with options and futures pricing models analysis. Section 5 presents the model choice that will lead to a VaR model for VIX futures along with results in section 6 followed by a discussion in section 7 on further improvements for the model based on the results as well as the conclusion in section 8. Please note that for sections 2, 3, and 4 which presents a literature review, notations will be preserved as in the original texts.
2 The VIX Index

2.1 The Origins

“It should be made clear at the outset that many indexes can be constructed to describe a certain economic phenomenon or to serve as a benchmark. Different indexes can serve different objectives or uses. Each index can be interpreted in such a way so as to answer a certain question. In principle, different objectives may call for different indexes in describing the time series of economic events. The search for a “universal” index, one which will satisfy all objectives, is futile. What should be looked for is an index which can satisfy well defined and well accepted objectives, one which can be used to better understand the economic phenomenon itself and its relationship to other economic phenomena. In order to achieve these objectives, the index numbers must be consistent over time. An index describes the history of a phenomenon. By construction, it is an ex-post measure. This fact emphasizes the constraints in using the index. One can try to use the index to predict the future by means of extrapolation, by using sophisticated statistical tools. But the index, by itself, does not tell what will happen next, only what has happened until the present” – Galai (1979)

Usual stock indexes, like the S&P 500, are calculated using rules that govern the selection of component securities. Using the prices of each security and the established set of rules, the indexes are then calculated as a value weighted portfolio of the component stocks. Instead of stocks, the VIX Index is rather comprised of options with the price of each option reflecting the market’s expectation of future volatility. Like other indexes, VIX employs a set of rules for selecting component options and a formula to calculate its value.

A decent amount of work has been involved in the creation and calculation methods of the VIX index. Given that it represents the options market on the S&P500, the creation of such an index would prove more difficult than an equity index due to the complexity of options. Since the 1970s, the literature presented ideas on the development of volatility indexes and derivatives whose payoff would be tied to them. Gastineau (1977) and Galai (1979) proposed option indexes similar in concept to stock indexes. Brenner and Galai (1989) proposed realized volatility indexes along with futures and options contracts on the indexes. Fleming & al. (1993) described the construction of the original VIX.

The first step came along with the listing of equity call options that began trading in 1973; Gastineau (1977) listed the requirements for a relevant option index to measure the premium levels on traded stock options. It is important to understand that an index which measures option premium levels must be
constructed differently from other common securities indexes since options are wasting assets, meaning the option’s value declines as time passes if the price of the underlying stock remains approximately unchanged. An option index should not measure the absolute but the relative magnitude of the premium over the option’s intrinsic value adjusted for variables such as dividends, interest rates, time to expiration, and the option’s strike to current stock price relationship. The author proposed two main methodologies, the first one being a value line index and the other an implied volatility index.

The value line index was constructed in the form of an average percentage premium with strike price and market price set equal and the expiration period standardized, interpolating from actual data when strike prices, market prices, and expiration periods vary. Two value line indexes were tested for three and six month options. While being a fair indicator of broad movements in premium levels at least for the six month index, the value line index did have a few drawbacks. The three month index seemed to be based on some options with much less than three months of life remaining, which caused certain instability when options were expiring. At times, the six month index has appeared to be overly affected by changes in the value option buyers attach to the possibility of getting long-term capital gain. Additionally, the method used in the construction of the value line indexes did not seem to compensate properly for changes in interest rates.

The second index that was suggested was based on implied volatility. The index provided a more precise measure of option premiums that could be related to changes in specific assumptions affecting the option’s values and to changes in the volatility of the underlying stock. The index itself measured the relationship between the volatility of the underlying stocks implied by the current level of option premiums and the historical volatility of underlying stocks. The author argued that an index relying on implied volatility was probably one of the simplest ways of determining ex-post whether a particular level of premium was too high or too low, given the subsequent volatility realized by the underlying stocks.

In the same line of thoughts, Galai (1979) made a proposal for two indexes on traded call options. He first described the implications of using indexes to explain certain economic phenomena and then applied the principle to a specific case of option index along with the possible objectives and limitations of such an index. The two suggested indexes were the following; one for a call buyer and another for a covered call writer. These two strategies were the most common for buying and selling options at the time. The index for the call buyer was in fact more volatile compared to the return on the underlying stock due to the high degree of leverage. The index for the covered call writer shows very mild changes since part of the risk of the short position in the option is canceled by holding a long position in the underlying stock. The two indexes were standardized with respect to time to maturity. Each index represented the return on 180 day average option. This procedure was done to maintain consistency with
respect to time. Only close to the money options were included as well. The author addressed the issue of varying trading interval created by non-trading days. Since the passage of time affects the value of options, non-trading days are included in today’s VIX.

Cox and Rubinstein (1985) refined the Gastineau procedure by including multiple call options on each stock and by weighting the volatilities in such a manner that the index is at the money and has a constant time to expiration.

A few years later, Brenner and Galai (1989) commented on the importance and the need of a volatility index for hedging purposes. They suggested three different realized volatility indexes, one for the equity market, one for the bond market, and a third one for the foreign currency market. They also argued and provided many examples that even if no volatility index can represent the exposure of all market participants and provide a perfect hedge for everyone; most potential users would find the instruments on a volatility index useful. They also provided a valuation method for volatility derivatives with the Cox et al. (1979) binomial model for option pricing.

With the launch of VIX in 1993, Fleming et al. (1995) provided a detailed description of how the index is constructed along with many statistical and econometrical proprieties. At its inception, the VIX rallied earlier efforts and extended the concepts on volatility indexes in two ways. First, it was and still is based on index options rather than stock options. While the “average” level of individual stock volatilities may be of passing interest, market participants are more concerned with portfolio risk or the level of risk after the idiosyncratic risks of the individual stocks have been diversified away. Second, it was based on the implied volatilities of both call and put options. This not only increases the amount of information incorporated into the index but also mitigates concerns regarding call/put option clienteles and possible discrepancies of the reported index level and the short-term interest rate. The first methodology for calculating the VIX used implied volatility of eight S&P 100 (OEX) options. The implied volatility was calculated using the Black and Scholes (1973) option valuation model for eight near-the-money (NTM) strikes (four calls and four puts) for the closest and second closest time to expiration.

During the latter part of the 1990s, the structure of index option trading in the U.S. changed in ways such as the most active index option market became the S&P 500 (SPX). Furthermore, the trading motives of market participants in market index options changed as well. In the early 1990s both index puts and index calls had balanced trading volume. Over the years, the index option market became dominated by portfolio insurers, who purchased out-of-the-money (OTM) and at-the-money (ATM) index puts for insurance purposes. Bollen and Whaley (2004) showed that the demand to buy OTM and ATM SPX puts is a key driver in the movement in implied volatility measures such as VIX.

A new methodology for the VIX calculation was then introduced on September 22, 2003. The new methodology now uses SPX options rather than OEX options.
and does not rely on the Black-Scholes model. It now relies on a more robust methodology for pricing continuously monitored variance swaps with a wider spectrum of out-of-the-money index calls and puts. These changes were made to better reflect the trading motives of market participants in market index options. Including additional option series would also help make the VIX less sensitive to single option price and therefore less susceptible to manipulation.

2.2 The History

Keeping in mind that the information an index offers is not necessarily about its current level but about the comparison of its level to its historical benchmark. A perfect example would be in the beginning of October of 2008 when the S&P 500 fell below the 950 mark. The relevance of that information at the time was not about the S&P 500 level, but rather about its level about a month earlier that was above 1250. Taking a look at the VIX history should be start in order to get a better understanding. Figure 1 shows historical data of the S&P 500 and the VIX from January 1990 to the end of May 2012.

![VIX and S&P500 Historical Data](image)

Figure 1: In blue is the time series of the SPX Index. In red the time series of the VIX Index

Perhaps the most interesting phenomenon that can be observed in the figure is that the VIX spikes upward at certain times and returns to more normal levels afterwards. The jumps in 1990 and 1991 correspond to the Gulf War. The first jump came in 1990, when Iraq invaded Kuwait and was then followed by a
second one in 1991, which corresponds to the United Nations forces attacks on Iraq. The next noticeable spikes are in October 1997 following a stock market sell-off in which the Dow-Jones fell 555 points and in October 1998 which was a period of general nervousness in the market. Then in October 2008, the VIX spiked to historical highs with the subprime crisis followed by several spikes later on due to the ongoing global financial crisis. Although the levels of the S&P500 and the VIX appear to spike in opposite directions, there are times when a run up in stock prices causes an increase in volatility. In January 1999, for example, the VIX was rising while the level of the S&P500 was rising. The same pattern can be seen in the first two months of 1995, June and July of 1997 and December 1999.

By looking at its history, it has become clearer why the VIX has been commonly referred to as the “Investor Fear Gauge” by the financial press. While volatility technically means dispersion of returns on either side of the mean, the S&P500 option market has been dominated by hedgers who buy index put options to protect their portfolios from a potential market drop. Since the put prices are driven up by this demand, the VIX could be seen as a price for portfolio insurance (Whaley (2009)). However, from a more theoretical standpoint, one can argue that the term “fear gauge” would clearly fit if the fear is of variance higher than expected. If the fear is of returns realizing lower than expected, then the suitability of the name may not be immediately obvious. The explanation provide by Carr and Lee (2009) is the following:

“Both the present VIX and the former VIX are constructed from OTM puts and OTM calls, with at-the-money (ATM) defined as the strike that equates call value to put value. According to the weighting scheme used in the CBOE white paper, the price of an OTM put receives more weight than the price of an equally OTM call, even if moneyness is measured by the log of the strike to the futures. This would seem to explain the fear gauge moniker except that in the standard Black model, an OTM put is cheaper than an equally OTM call. However, a detailed calculation shows that if the Black model is holding and a continuum of strikes are available, then more dollars are invested in OTM puts than in OTM calls when replicating a continuously monitored variance swap (with no cap). This calculation also shows that this bias toward OTM puts is purely due to the convention that the ATM strike is the forward price. If the ATM strike is redefined to be the barrier that equates the Black model theoretical values of an up variance swap to a down variance swap, then OTM puts cost as much in the Black model as OTM calls. By this metric, positive and negative returns on the underlying are weighted equally in the Black model. The empirical reality that an OTM put typically has a higher implied volatility than an equally OTM call motivates that the fear in the term “fear gauge” is of returns realizing below the benchmark set by the alternative ATM strike.”
Additionally, one has to keep in mind that the VIX Index is a forward looking index. It measures the volatility the investors expect to see. In other words, it is a market consensus on the expected volatility for the upcoming 30 days. This provides a certain explanation to the fact that at a number of times in its history the VIX was above its average even though the S&P 500 seemed to be doing well. The technology bubble at the end of the 1990’s is a perfect example. While at the time nothing seemed to stop the bullish market that was taking place, a general feeling of nervousness from the investors could be felt has to when the exceedingly over-valued technology stocks would return to their fair value. As history shows, the volatility was higher than usual starting from 1998 and going into the new millennium even in the bullish market. This shows that at that point in time, the investors were expecting to see the bullish market become bearish sooner than later.

2.3 The Methodology

Like any other index, the VIX follows a precise methodology (CBOE, 2003) in order to derive its price. As section 2.1 reveals, its methodology as evolved over the years to better reflect the market’s reality. Initially derived from the average implied volatility of four NTM calls and four NTM puts using the models Black-Scholes, the methodology has evolved in order to include all OTM calls and puts and to not rely on the Black-Scholes model anymore.

More precisely, the VIX is a continuously monitored variance swap. It is calculated as the square root of the 30 days average variance swap rate using the near- and next term SPX options. To minimize pricing anomalies, near term options must have at least one week to expiration. When the near term options have less than a week to expiration, the second and third months are used.

Variance or volatility swaps are conceptually similar to any other conventional swaps. A variance swap is an OTC contract with zero upfront premium but unlike most, it only has payment at expiry. An interest rate swap, for example, which exchanges fixed interest payments against floating interest payments on a predetermined notional has in most cases periodic payment dates until expiry. In the case of a variance swap, the payment is made only at expiry and is settled as follows.

The long side of the variance swap pays a positive dollar amount base on the variance swap rate agreed upon at inception. In return for this fixed payment, the long side receives a dollar amount at expiry, called the realized variance of the underlying index. Conventionally, the realized variance is an annualized average of the squared daily returns. If the realized variance is higher than the rate agreed upon at inception, the investor will gain the difference between the realized amount and the swap rate. If the realized variance is lower than the rate, the investor will take a loss equal to the difference between the realized amount and the variance swap rate.
As with interest rates swaps, both counterparties will agree upon a fair-value for the fixed rate at inception. The variance swap price can be calculated with the following equation.

\[
V_0^2(T) = -2 \int_0^1 \frac{1}{K^2} \text{Put}(K,T) \, dK - 2 \int_1^\infty \frac{1}{K^2} \text{Call}(K,T) \, dK
\]  \tag{1}

However, for equation (1) to hold, a continuum of strikes must be available. To solve the problem, the CBOE uses a replication method to approximate the equation with a discreet set of available strikes.

\[
\sigma_T^2 = \frac{2}{T} \sum \frac{\Delta K}{K_i^2} e^{RT} Q(K_i) - \frac{1}{T} \left[ \frac{F}{K_0} - 1 \right]^2
\]  \tag{2}

Where:

- \( \sigma^2 \): Price of the variance swap
- \( T \): Time to expiration in years
- \( F \): Forward index level derived from index option prices
- \( K_0 \): The first strike below the forward index level \( F \)
- \( K_i \): Strike price of the \( i \)th OTM option;
  - a call if \( K_i > K_0 \) and a put if \( K_i < K_0 \); both put and call if \( K_i = K_0 \)
- \( \Delta K_i \): Interval between strike prices; half the difference between the strikes on either sides of \( K_i \); \( \Delta K_i = \frac{K_i + 1 - K_i - 1}{2} \)
- \( R \): Risk Free interest rate to expiration
- \( Q(K_i) \): The mid-point of the bid-ask spread for each option with strike \( K_i \)

Note: For the lowest strike \( \Delta K \) is simply the difference between the lowest strike and the next higher strike. Likewise, \( K \) for the highest strike is the difference between the highest strike and the next lower strike.

After finding the variance swap price for the near and next term SPX options, the VIX can be calculated and using equation (3).

\[
VIX = 100 \times \sqrt{\left\{ T_1 \sigma_1^2 \left[ \frac{N_{T_2} - N_{30}}{N_{T_2} - N_{T_1}} \right] + T_2 \sigma_2^2 \left[ \frac{N_{30} - N_{T_1}}{N_{T_2} - N_{T_1}} \right] \right\} \times \frac{N_{365}}{N_{30}}}
\]  \tag{3}

Where:

- \( N_{T_1} \): number of minutes to settlement of the near-term option
- \( N_{T_2} \): number of minutes to settlement of the next-term option
- \( N_{30} \): number of minutes in 30 days = 43,200
- \( N_{365} \): number of minutes in a year = 525,600
2.4 The Purpose

The VIX index may serve purpose to survey the market expectancy on upcoming volatility, but it also serves purpose to investor’s who wish to transfer (hedge) their volatility risk. With the increasing popularity of over-the-counter volatility derivatives in the 1990’s, many argued the growing need of simple and standardized volatility hedging tools. The arrival of an index such as the VIX would provide a benchmark on which those contracts could be written. Even more so, the VIX derivatives who made their appearance in the early 2000’s can now be used directly to hedge volatility risk.

The VIX derivatives provide a cost-effective way for hedging volatility risk for portfolio insurers, options market-makers and covered call writers. Whaley (1993) shows that the cost effectiveness to hedge for market volatility risks is enhanced by using VIX derivatives as compared to hedging with index options.

To demonstrate the cost effectiveness of hedging with volatility derivatives, lets take a look at the simplified example from Whaley (1993).

Consider a trader whose portfolio contains options or securities with option-like features. The two most important risk factors are a change in the underlying security price and a change in the expected volatility. Those risks are measured by computing the delta and vega of the option portfolio. Keeping in mind that delta is the level of change in the option or portfolio price with respect to the level of change of the underlying price and vega the level of change in the option or portfolio price with respect to the level of change in the underlyings volatility.

Assuming the option values obeys the Merton (1973) constant proportional dividend yield option valuation formula and the prices of the volatility options are assumed to obey the Black (1976) futures option valuation formula, the price of a European call option on a stock index is given by the following,

\[ c = e^{-δt}SN(d1) - Ke^{-rT}N(d2) \]  

Where \[ d1 = \frac{ln(S/K) + (r - δ + 0.5σ^2)T}{σ\sqrt{T}} \]

And \[ d2 = d1 - σ\sqrt{T} \]
Where $K$ is the exercise price, $T$ is the time to maturity, $N$ is the cumulative normal probability, $S$ is the price of the underlying, $\delta$ is the dividend yield, $\sigma$ the underlying volatility and $r$ is the risk-free rate.

The effect of change in the underlying price on European call options value is delta and is measured as follows,

$$Delta_c = e^{-\delta T} N(d_1) > 0 \quad (5)$$

For a European put option on a stock index, the price is given by,

$$p = Ke^{-rT} N(-d_2) - e^{-\delta t} S N(-d_1) \quad (6)$$

And delta is,

$$Delta_c = e^{-\delta T} N(-d_1) < 0 \quad (7)$$

For risk management purposes, we focus on the overall portfolio delta which is the sum of all individual deltas times the number of contract held.

$$Net \ Portfolio \ Delta = \sum_i^n \Delta_i \times Number \ of \ Contracts_i \quad (8)$$

Where $n$ is the number of different contract in the portfolio.

The effect of change in volatility on the option value is measured by the option’s vega. Using the equation above to value index calls and puts, the vega of both call and put options are equal and is given by,

$$Vega_c = Vega_p = Se^{-\delta T} n(d_1) \sqrt{T} > 0 \quad (9)$$

Where $n(d_1)$ is the normal density function evaluated at $d_1$.

From this equation, it can be deduced that an increase in volatility will correspond to an increase in both index calls and puts values since there is greater probability of a large index move during the life of the option. Like the net delta, the net vega of an index option portfolio is the sum of the weighted volatility exposures of the individual option series.
Let us now consider an index option market maker who acquires a large option position during his day and holds it overnight. The portfolio is presented in panel A of table 1. It can be deduced that the position is short market volatility since the market maker is short on all four options. The net vega of the portfolio is -196.700. This implies that, if volatility increases by 100 basis point overnight, the option portfolio value will decrease by about $197 or a little more than 4%. We can now compute the price risk exposure of the portfolio, the delta, which is 0.025. The portfolio is said to be delta neutral. If the index level rises by a dollar, the portfolio will increase only by 2 cents.

There are now four hedging scenarios to consider which are presented in panels B, C, D, and E of table 1. The first two scenarios involve hedging with index options. Initially using only put options (Panel B) and then using both put and call options (Panel C). The last two scenarios involve hedging with volatility derivatives. The first one is hedging with volatility futures (Panel D) and the last with volatility options (Panel E).

Let’s take a look at the first scenario (Panel B). To counter his volatility risk, the market maker can buy either index puts or index calls, since both will increase in value with volatility. Since the portfolio’s current net vega is 196.700, the net vega of the purchased index option must be equal 196.700.

Suppose that an index put with an exercise price of 395 and a time to expiration of thirty days is available. Its price, delta, and vega are shown in Panel B. To eliminate volatility exposure using the 395 put, the market maker must then buy 448 contracts

\[ n_p = \frac{196.70}{0.439} \approx 448 \]

The cost of the options for hedging is $2920.96. With this hedge, the portfolio is now vega neutral with a net vega of -0.028.

Having neutralised the vega exposure by buying an index put option, the market maker has however created another problem. He has changed his delta exposure. After the purchase of the put options the net portfolio delta has now become -174.695. This means that an increase of one point overnight will drop the portfolio value by
nearly 11%, a situation that would most likely cause the market maker to lose sleep that night.

To avoid changing the delta risk while hedging for volatility, at least two index options must be used. A common way to hedge this risk is to use a volatility spread, which consist in buying a call and a put in such way that the net delta value is zero. Assume now that the market maker buys a 405 call with thirty days to expiration in addition of his 395 put with thirty days as shown in Panel C.

To find the optimal number of calls and puts to buy, we now solve a simultaneous system of equations. Since the current delta of the portfolio is already approximately zero, we want the net delta of the newly purchased options to be zero. To hedge the vega exposure, we want the net vega of the newly purchased options to be 196.700.

\[
\begin{align*}
    n_p \Delta p + n_c \Delta c &= 0 \\
    n_p \text{Vega}_p + n_c \text{Vega}_c &= 196.700
\end{align*}
\]

By solving this system of two equations with two unknowns, we find \(n_p = 233\) and \(n_c = 209\) for a total cost of $3019.78. After the hedge is in place the net delta exposure is 0.279 and a net vega exposure of -0.154, which means the portfolio is now delta and vega neutral. The effective cost of using this hedge is the erosion of the options value overnight (Theta).

The theta of an option is the change in the options value with respect to time to maturity. The theta values of the 395 put and 405 call are shown in Panel C. The net theta on the purchased index options is 64.710, which means that holding everything else equal the hedge will lose about $64.71 overnight.

To determine the effectiveness of the volatility hedge, we compare the profit and losses (P&L) for the unhedged portfolio to the P&L of the hedged portfolio when the volatility rate changes unexpectedly overnight.

Figure 2 shows that the unhedged portfolio has considerable volatility exposure. If the volatility rate increases increases from 20% at the close of trading to 24% by the following morning, the portfolio value loses a little less than $800 which is about 18% of the portfolio value.

The hedged portfolio, however, is relatively immune to shifts in the volatility rate. Overnight shifts in the volatility rate as high as 500 basis points in either direction have marginal effect on portfolio value.
Hedging the vega exposure using volatility futures is straightforward. Since the net vega of the market maker’s portfolio is -196.700, the portfolio value will decrease by $196.00 for each 100 basis points of volatility increase. Since the price of volatility futures moves directly with volatility, a 100 basis point increase in the volatility represents a $1.00 increase in the futures price. Therefore, the optimal number of volatility futures to buy is 197.

Panel D of table 1 shows the net effect of buying 197 volatility futures contracts. Since there is no cost in establishing the futures position, the portfolio value remains at $4516.25. After the futures are purchased, the net vega of the portfolio is reduced to 0.300. The net delta of the portfolio remains unchanged. Volatility derivatives, unlike index options, do not affect the delta exposure of the index option portfolio.

Figure 3 shows the P&L of the unhedged portfolio with the P&L of the hedged portfolio including volatility futures against changes in volatility level. As shown in the figure the hedged portfolio’s P&L remains fairly constant over the considered volatility range.

Hedging the vega risk with volatility options is nearly as straightforward as using volatility futures. The only difference is that the value of a volatility option does not move as quickly as volatility futures in response to changes in volatility level. To quantify the rate of
<table>
<thead>
<tr>
<th>Table 1: Volatility Hedging Strategies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quantity</td>
</tr>
<tr>
<td>----------</td>
</tr>
<tr>
<td>A. Unhedged Portfolio</td>
</tr>
<tr>
<td>-50</td>
</tr>
<tr>
<td>-100</td>
</tr>
<tr>
<td>-75</td>
</tr>
<tr>
<td>-100</td>
</tr>
<tr>
<td>B. Hedged Portfolio Using Put Options Only</td>
</tr>
<tr>
<td>Total Hedge Position</td>
</tr>
<tr>
<td>Net Portfolio Position with Hedge</td>
</tr>
<tr>
<td>C. Hedged Portfolio Using Index Call and Put Option</td>
</tr>
<tr>
<td>Total Hedge Position</td>
</tr>
<tr>
<td>Net Portfolio Position with Hedge</td>
</tr>
<tr>
<td>D. Hedged Portfolio Using Volatility Futures</td>
</tr>
<tr>
<td>Total Hedge Position</td>
</tr>
<tr>
<td>Net Portfolio Position with Hedge</td>
</tr>
<tr>
<td>E. Hedged Portfolio Using Volatility Call</td>
</tr>
<tr>
<td>Total Hedge Position</td>
</tr>
<tr>
<td>Net Portfolio Position with Hedge</td>
</tr>
</tbody>
</table>

Notes: The index portfolio’s level is assumed to be 400; its volatility rate is 20%, and its dividend yield 3%. The volatility index level is assumed to be 20%, and the volatility rate of the volatility index is 75%. The interest rate is 5%.

To change, we need to compute the volatility option’s delta. To do so, we apply the Black (1976) futures option valuation framework. The valuation for a European-style call option on the volatility index is

\[
c = e^{-rt}[F_vN(d1) - K N(d2)]
\]  

Where\[
d_1 = \frac{ln(S/K) + (0.5\sigma_v^2)T}{\sigma_v \sqrt{T}}
\]

And

\[
d_2 = d_1 - \sigma_v \sqrt{T}
\]

The delta value of a volatility call option is
\[ \Delta_c = e^{-rT}N(d1) > 0 \]  

(13)

The valuation of a European style put option is

\[ p = e^{-rt}[KN(d2) - F_vN(-d1)] \]  

(14)

With a delta value of

\[ \Delta_p = -e^{-rT}N(-d1) < 0 \]  

(15)

To set the volatility hedge for the market maker’s portfolio, we simply divide the index option portfolio’s vega exposure by the volatility option’s delta, since the delta for a volatility option measures the option’s sensitivity to volatility change.

In the example, the unhedged portfolio has a net vega of -196.700. Suppose the market maker has the opportunity to buy volatility index calls with an exercise price of 20 and a time to expiration of thirty days. The optimal number of volatility calls to buy is

\[ n_{cv} = \frac{196.700}{0.541} \approx 36 \]

Figure 3: Unhedged P&L (dotted line) and hedged P&L with volatility futures (dashed line) and volatility options (solid line)
Panel E of table 1 summarizes this hedge. The total of purchasing volatility calls is $622.24. With the purchase of the calls, the net portfolio vega is reduced to 0.224. The net delta remains unchanged once again since volatility derivatives do not affect delta exposure.

To measure the effectiveness of the volatility option hedge relative to the volatility futures hedge, let’s look at figure 3. While the volatility futures hedge is immune to large volatility shifts, the volatility option hedge provides an interesting convexity in which the market marker’s profits increase in the event of a volatility shift overnight. However, this convexity comes with a cost since the volatility options are expected to lose about $10.19 in value overnight.

As seen in the example above, using volatility derivatives to hedge market volatility offers at least two advantages over index option contracts. We not only find the hedge simpler to implement since only one volatility derivative contract is required instead of two index option contracts but also, volatility derivatives are cheaper. Hedging with volatility futures incurs no costs and hedging with volatility options will cost only a fraction of a similar hedge using index options. Even if this simple example is making a number of simplifying assumptions regarding both index and volatility options, it still shows the mechanics and advantages of hedging with volatility derivatives.

3 VIX ETPs

In recent years, volatility trading has become the new trend. Many volatility exchange traded products (ETPs) linked to the VIX Index have attracted many traders. More than 30 VIX ETPs are actually listed and are seeing their fair amount of success with an aggregated market investment value of nearly $4 billion and trading volume that averages $800 millions daily.

As attractive as these products may seem for investors, many will find themselves victim of a far from pleasant surprise. The idea of holding a security that is negatively correlated with the market seems promising for diversifying a portfolio. It also seems that many investors think they are directly holding the VIX Index in their portfolio when buying these volatility ETPs but they are bound to find out that it is sadly not the case.

Unlike most exchange traded securities, the VIX ETPs not suitable for buy-and-hold investments and are guaranteed to lose money. These ETPs are in fact created from VIX futures trading strategies that demand daily rebalancing and are not only subject to management fees and expenses, including futures commissions and trading fees, licensing fees and forgone interest income but are also bound to lose money from a contango trap which makes the VIX futures constantly drawn downward towards the level of the VIX Index.
A recent article from Robert Whaley (Whaley, 2013), discusses the motives of volatility ETPs holders and also demonstrate the ineffectiveness of holding such a security in a long-term portfolio. It seems that most investors that hold this type of security are uninformed and may think for the most part that holding a VIX ETP will diversify their portfolio. However, it is clearly shown that the losses incurred over the holding period will wash away completely any diversification effect.

Although, the VIX ETPs seem to be effective in tracking their respective indexes, the problem remains in the VIX futures index being tracked. These indexes are intended to mimic the behaviour of futures trading strategies that involve rolling a VIX futures position in a manner that maintains a constant maturity.

The problem with this kind of strategy becomes obvious when looking at the VIX futures curve. The contango effect mentioned earlier can be seen in figure 4 where the slope of the curve is consistently upwards. Suppose the curve remains the same shape throughout time, then with every passing day the dots on the curve (the futures prices) should be marching down the curve as the VIX futures expiration grow short to ultimately converge to the current VIX value.\(^3\) In other words, the VIX has a constant maturity of 30 days and will remain in place as time passes by. However, the maturity of the VIX futures does not stay constant since they are set to expire on a determined date.

![VIX Futures Price Curve on September 23, 2014](image)

Figure 4: VIX Futures Curve

\(^3\)Also referred to as roll-down.
The VIX futures curve being more often than not in contango will cause a long position in the futures to systematically lose money. For the strategies employed with tracking indexes, the losses incurred are predictable by taking the slope straddling the constant maturity. For example, for a 30 days constant maturity with daily rebalancing, the slope between the first and the second maturity is 

\[(15.54 - 14.90)/20 = 0.032,\]

holding everything else equal, this position stands to lose $0.032 or 0.21% on every passing business days which yields whopping 52.9% loss annually.

### 3.1 Specific Case: Short-Term Volatility Tracking Index ETN (VXX)

In this case we will concentrate on the VXX which is one of the most traded volatility ETP. Since it was the first VIX ETP launched along with its mid-term volatility companion the VXZ, they both enjoyed a first-mover advantage. As of March 30 2012, the VXX had a market capitalization of $1.86 billion and traded more than 32 million shares a day going from January through March 2012, which is about the same in dollar volume as Ford shares.

The VXX tracks the VIX ST ER futures index that models returns from a long VIX futures position that is rolled continuously throughout the period between first two contracts expiration dates. From the iPath prospectus, here is how the index is constructed.

The total return version of the ETN incorporates interest accrual on the return of the notional value and reinvestment of returns and interest. Interest accrus based on the 3-month U.S. Treasury rate. The S&P 500 VIX Short-Term Futures Index measures the return from a rolling long position in the first and second month VIX futures contracts. The Index rolls continuously throughout each month from the first month VIX futures contract into the second month contract as shown in more detail below.

On any S&P 500 VIX Futures business day, \( t \), the short-term volatility index is calculated as follows:

\[ \text{IndexTR}_t = \text{IndexTR}_{t-1} \times (1 + \text{CDR}_t + \text{TBR}_t) \]  

Where \( \text{IndexTR}_{t-1} \) is the index tracker on the preceeding business day and \( \text{CDR}_t \) is the Contract Daily Return, given by

\[ \text{CDR}_t = \frac{\text{TDWO}_t}{\text{TDWO}_{t-1}} - 1 \]

Where \( \text{TDWO}_t \) is the Total Dollar Weight Obtained on \( t \) as given by,

\[ \text{TWDO}_t = \sum_{i=1}^{2} \text{CRW}_{i,t-1} \times \text{DCRP}_{i,t} \]
And

$$TWDO_{t-1} = \sum_{i=1}^{2} CRW_{i,t-1} \times DCRP_{i,t-1}$$ (19)

Where $CRW_{i,t}$ is the Contract Roll Weight of the $i^{th}$ VIX futures contract on date $t$ and $DCRP_{i,t}$ is the Daily Contract Reference Price of the $i^{th}$ VIX futures contract on date $t$.

The Treasury Bill Return $TBR_t$ is calculated with the following formula,

$$TBR_t = \left[ \frac{1}{1 - \frac{91}{360} \times TBAR_{t-1}} \right]^{\frac{\Delta t_{t-1}}{91}} - 1$$ (20)

Where $\Delta t_t$ is the number of calendar days between the current and previous business days and $TBAR_{t-1}$ is the most recent weekly high discount rate for 91-day US Treasury bills\(^4\) effective on preceding business day. Generally the rates are announced by the US Treasury on each Monday. On Mondays that are bank holidays, Fridays rate will apply.

The Roll Period starts on the Tuesday prior to the monthly CBOE VIX Futures Settlement Date (the Wednesday falling 30 calendar days before the S&P 500 option expiration for the following month), and runs through the Tuesday prior to the subsequent month's CBOE VIX Futures Settlement Date. Thus, the index is rolling on a continual basis. On the business date after the current roll period ends the following roll period begins. In calculating the total return of the index, the Contract Roll Weights of each of the contracts in the index, on a given day, $t$, are determined as follows,

$$CRW_{1,t} = 100 \times \frac{dr}{dt}$$

$$CRW_{2,t} = 100 \times \frac{dt - dr}{dt}$$ (21)

Where:

$dt$ is The total number of business days in the current Roll Period beginning with and including, the starting CBOE VIX Futures Settlement Date and ending with, but excluding, the following CBOE VIX Futures Settlement Date. The number of business days stays constant in cases of a new holiday introduced intra-month or an unscheduled market closure.

$dr$ is the total number of business days within a roll period beginning with, and including the following business day and ending with, but excluding, the following CBOE VIX Futures Settlement Date. The number of business days includes a new holiday introduced intra-month up to the business day proceeding such a holiday.

\(^4\)Bloomberg ticker: USB3MTA.
At the close on the Tuesday, corresponding to the start of the roll period, all of the weight is allocated to the first month contract. Then on each subsequent business day a fraction of the first month VIX futures holding is sold and an equal notional amount of the second month VIX futures is bought. The fraction, or quantity, is proportional to the number of first month VIX futures contracts as of the previous index roll day, and inversely proportional to the length of the current roll period. In this way the initial position in the first month contract is progressively moved to the second month contract over the course of the month, until the following roll period starts when the old second month VIX futures contract becomes the new first month VIX futures contract.

In addition to the transactions described above, the weight of each index component is also adjusted every day to ensure that the change in total dollar exposure for the index should only be due to the price change of each contract and not due to using a different weight for a contract trading at a higher price.

A glance at figure 5 allows to see that the VXX tracks the SPVXSTR Index decently when putting both on the same scale. As explained earlier, the VXX consistent losses are not due to its inability to track the index but because of the roll-down effect of VIX futures that are drawn towards the spot VIX. During the presented period from the VXX inception on January 29th 2009 going to June 13th 2012, it has lost 94.85% of its value. Meanwhile it has also been subjected to a 1 for 4 stock split on September 11th 2010 and subsequently two other 1 for 4 splits later on in 2012 and 2013. Figure 6 shows the slope between the second and first VIX contract during the period which averages $-0.062$/day. Figure 7 shows the slope during the period in percentage which averages $-0.288\%$ daily ($-72\%$ annually).

![Comparison between SPVXSTR and VXX since VXX inception on 2009/01/29](image.png)
Referring to the previous evidence, it becomes clear that the slope between the VIX futures has a dominant role in determining the daily return of the contracts or any VIX ETPs. To back the evidence, Whaley (2013) also shows with a regressive model that the slope between the futures contract as a statistically significant role in forecasting the VIX ETPs returns. These facts should shed light as to which model should be suitable to model the VIX futures and subsequently the VIX ETPs in a manner to capture the roll-down effect of the futures.
4 Volatility Derivatives

Volatility derivatives\(^5\) were first traded at the beginning of the 1990’s on the OTC market and rapidly became popular amongst traders, especially variance swaps.

Following this buildup of OTC activity, the CBOE introduced in 1993 its first volatility index, the VIX. CBOE introduced the index to provide a benchmark of expected short-term market volatility but also an index upon which futures and options contracts on volatility could be written. However, it is only after the release of the new VIX in 2003 that the Chicago Futures Exchange (CFE) launched, in March 2004, the exchange traded VIX futures. Following the success of the VIX futures, the CFE introduced, in February 2006, VIX options. Like the futures, the VIX options mature on the one day of each month when only a single maturity is used to compute the VIX. VIX options are among CBOEs most liquid contract, second after the SPX index options. Their popularity comes from the wider spectrum of possibilities they offer. Since the SPX and VIX are highly negatively correlated, one can easily compare a call option on the VIX as a put option on the SPX.

Meanwhile, the research started to address a new problem as to how to evaluate volatility derivatives. The theoretical effectiveness of hedging with volatility derivatives had been demonstrated in Whaley (1993) but the solution to pricing volatility futures was still problematic (and still is to this day). Volatility futures have a number interesting properties that differentiate them form other futures contracts such as commodities or equity futures and differ from the standard cost-of carry model in a number of ways.

The volatility derivatives literature quickly followed the VIX inception in 1993. Even more so after the CBOE launched standardized VIX futures contracts in 2004 along with VIX options in 2006 and since, many authors have addressed the issues of VIX derivatives pricing in a variety of frameworks.

The first VIX option pricing model used was in fact the very simplistic Black (1976) futures option pricing formula. Since a number of models has been presented within the equilibrium model framework. The scope of this framework is mainly to explain the VIX futures curve and attempt to price the VIX futures from the instantaneous variance of the S&P500 using an analytical formula that can usually be solved into a closed-form option pricing formula. Some work has also been done within another framework usually referred to as the No-Arbitrage framework\(^6\) which consists in taking the volatility term-structure for granted and using it to price more complex derivatives. This approach does not try to explain either the term-structure nor option prices observed in the

\(^5\)The terms “Volatility” and “Variance” may be used without distinction. The first one being the square root of the other. Volatility will be used throughout the text unless specifications have to be made.

\(^6\)The no-arbitrage framework as been commonly used in the pricing of interest rates derivatives.
market but rather uses these traded assets to shape higher order assets such as path-dependant claims or any other types of exotic derivatives.

4.1 Black-Type models

The Black-Type models are surely the most simplistic way for pricing VIX options from VIX futures prices. The two main models in this category are Whaley (1993) and Carr and Lee (2007).

4.1.1 Whaley (1993)

The reference to the Whaley (1993) model in the VIX literature refers in fact to the Black (1976) futures option pricing formula. Assuming log-normality of the VIX futures, the price of a VIX call is given by equation (4), and the put price by equation (6).

4.1.2 Carr & Lee (2007)

In the same framework as Carr and Wu (2005) that presented a lower and upper bound of the VIX futures price using theory on variance and volatility swaps by applying Jensens inequality. Showing that the current VIX futures (with maturity $T$) price is between the forward volatility swap rate and the forward variance swap rate over the period of $(T, T + 30/365)$, Carr and Lee (2007) suggested a new model-free approach. Unlike the preceding models this model is not model-dependant. It rather relies on market quotes of variance and volatility swaps as inputs. There is no need for calibration from historical VIX options data and makes real-time pricing and implementation of hedging strategies readily applicable. The VIX options valuation formula in this case is an application of their more generalized pricing formula for options on realized volatility.

Variance swaps prices depend on the expectation and volatility of variance. The expectation is revealed by the swap price itself, and the volatility can be inferred from variance and volatility swaps prices together which in turn enables to valuate volatility or variance options.

The argument they provide is the following,

Let $S_t$ denote the value of the stock price at time $t$. Defining $R^2$ as a continuously monitored variance swap, meaning the quadratic variation of $\log(S)$ times a constant conversion factor $u^2$ that includes any annualization or rescaling we have the following equation,
\[ R_{r,t}^2 = u^2 \sum_{\tau < t_n \leq t} \left( \log \frac{S_{t_n}}{S_{t_{n-1}}} \right)^2 \]

Let \( A_t \) be the time-\( t \) value of the variance swap which pays \( R_{0,T}^2 \)

Let \( B_t \) be the time-\( t \) value of the variance swap which pays \( R_{0,T} \)

Let \( r \) be the assumed constant interest rate, and let \( A_t^* = A_t e^{r(T-t)} \) and \( B_t^* = B_t e^{r(T-t)} \) be the time-\( t \) variance swap rate and volatility swap rate respectively.

The volatility swap’s concave square-root payoff is dominated by the linear payoff consisting of \( \sqrt{A_t^*} \) in cash, plus \( 1/(2\sqrt{A_t^*}) \) variance swaps with fixed leg \( A_t^* \). The denomerating payoff has forward value \( \sqrt{A_t^*} \), because the variance swap value is zero. This enforces Jensen’s inequality \( \sqrt{A_t^*} \geq B_t^* \) by superreplication.

This concavity’s price impact, measured by how much \( \sqrt{A_t^*} \) exceeds \( B_t^* \) depends on the volatility of the volatility with the following relation.

\[ A_t^* - (B_t^*)^2 = E_t[R_{0,T}^2] - E_t[R_{0,T}]^2 = Var_t[R_{0,T}] \] (23)

The volatility of volatility can then be found by obtaining the swap value \( A \) and \( B \), which allows to price options on \( R_{0,T}^2 \) and \( R_{0,T} \).

B applying a special case of their generalized formula by using the time-0 SPX implied variance swap \( A_t \) and the VIX for volatility swap \( B_T \), call options can be valued with the following equation,

\[
C(B_t, K, T) = e^{-rt}(B_t N(d1) - KN(d2))
\]

\[
\mu_1 = B_t \\
\mu_2 = A_t e^{rt} \\
m_t = 2\ln(\mu_1) - 0.5\ln(\mu_2) \\
s_t^2 = \ln(\mu_2) - 2\ln(\mu_1) \\
d_1 = \frac{m_t - \ln(K)}{s_t} + s_t \\
d_2 = d_1 - s_t
\]

Where \( B_t \) is the volatility swap rate, \( A_t \) is the variance swap rate, \( m_t \) is the time-\( t \) conditional mean and \( s_t \) is the variance of the log return of the VIX.

Being also a Black-Type model, the difference between the Carr and Lee model and the Whaley model is that there is no need to estimate the constant volatility \( \sigma \) since we compute the forward volatility \( s_t \) of the underlying VIX from available
market data. However the model does not account for correlations different than 1 between forward variance maturities.

4.2 Equilibrium Models

The literature on equilibrium models for the VIX futures started soon after the VIX inception. Grünbichler and Longstaff (1996) proposed a model similar to the Heston (1993) model that would directly model the dynamics of the VIX term structure. They also provided closed form formulas to price the VIX futures and options. After the VIX methodology changed in 2003 to become the square-root of a 30 days variance swap on the S&P500, Zhang and Zhu (2006) tried modeling the dynamics of the VIX futures directly within the Heston framework with the argument that modeling the instantaneous variance of the S&P500 should allow to price the VIX futures. A few years later Sepp (2008) as well as Lin and Chang (2009) brought a more sophisticated approach of the Heston model to model the VIX futures by adding a jump components in the volatility and the underlying’s processes.

4.2.1 Grünbichler & Longstaff (1996)

Grünbichler & Longstaff took evidence from the empirical literature that shows the presence of auto-correlation and behavior of mean-reversion in the volatility process which comes against the assumption of independent log-normality. They argued that these features have significant implications on the hedging behavior and pricing of volatility derivatives.

At the time, there were no active VIX futures or options contracts being traded except for over-the-counter deals. Since there were no exchange quotes for the VIX futures and knowing the standard cost-of-carry model could not apply, this would cause valuing options with the Black (1976) model to be problematic. They however presented a closed-form formula using a square root stochastic volatility process similar to Heston (1993)\textsuperscript{7} based on the following principles.

- The volatility futures prices are bounded above zero which means the volatility futures price does not converge to zero when volatility tends towards zero. If the current volatility reaches zero, it should immediately return to a positive value. Thus, the expected value of the volatility futures is strictly greater than zero even when the current value of the volatility is zero.

- The basis can also be either positive or negative, meaning that in a risk-free environment where all securities must earn the riskless rate of return, 

\textsuperscript{7}Grünbichler and Longstaff (1996) directly model the process of the VIX as compared to Heston (1993) which models both the index and its volatility, therefore modeling the VIX indirectly in this case.
the expected return on volatility can both be positive or negative and will generally not equal the riskless rate which makes the standard hedging arguments not applicable.

• The longer term futures should converge to the long run mean of volatility. This implies that longer-term futures contracts may not be effective instruments for hedging volatility risk. Due its mean reversion, any change in the current volatility is expected to be reversed prior to the expiration of the contract. This long-run mean convergence also affects the pricing and hedging behavior of the volatility options.

From these notions come the dynamics of the volatility $V$ given by

$$dV = (\alpha - \kappa V)dt + \sigma \sqrt{V}dW$$

Where $\alpha$, $\kappa$ and $\sigma$ are constants and $W$ is a standard Wiener process.

The process captures the features of a mean-reverting AR(1) process and also allows the variance of changes in implied volatility not to be constant, but to increase with the level of volatility. Given empirical estimates of the mean, variance and serial correlation, the parameters $\alpha$, $\kappa$ and $\sigma$ can be easily obtained by inverting the analytical expressions for the mean, variance and serial correlation since the mean of the stationary distribution corresponds to $\alpha/\kappa$, the variance to $\alpha \sigma^2 / 2\kappa^2$, and the serial correlation to $e^{-\kappa \Delta t}$.

Let’s now consider a contingent claim with payoff $B(V_T)$ at time $T$ depending only on $V_T$. Knowing that $V$ is not the price of a traded asset, this allows for the possibility that volatility risk is priced by the market. Being consistent with Wiggins (1987), Stein and Stein (1991), and others, Grünbichler and Longstaff (1996) make the assumption that the expected premium for volatility is proportional to the volatility level, $\zeta V$. This assumption is similar to the implications of general equilibrium models such as Cox et al. (1985).

Given this framework, the current value of the claim, $A(V,T)$, satisfies the fundamental valuation equation

$$\frac{\sigma^2}{2}VA_{VV} + (\alpha - \beta V)A_V - rA = A_T$$

Where $\beta = \zeta + \kappa$, subject to the expiration date condition

$$A(V_T, 0) = B(V_T)$$

Let $D(T)$ denote the current price of a $T$-maturity riskless unit discount bond. The solution to this partial differential equation can be expressed as

$$A(V_T, 0) = D(T) E \left[ B(V_T) \right]$$
Where the expectation is taken in respect to the risk-adjusted process for $V$

$$dV = (\alpha - \kappa V)dt + \sigma \sqrt{V}dW$$ \hfill (29)

This risk-adjusted process implies that $\gamma V_T$ is distributed as a non-central chi-squared variate with $\nu$ degrees of freedom and non-centrality parameter $\lambda$, where

$$\gamma = \frac{4\beta}{\sigma^2(1 - e^{-\beta T})}$$ \hfill (30)

$$\nu = \frac{4\alpha}{\sigma^2}$$

$$\lambda = \gamma V e^{-\beta T}$$

Following this, the futures prices can be derived. Let $F(V, T)$ denote the futures price for a contract on $V$ with maturity $T$ that can be expressed as the expected value of $V$ at time $T$.

$$F(V, T) = \mathbb{E}[V_T]$$ \hfill (31)

Where the expectation is taken with respect to the risk-adjusted process for $V$ from equation (29). Evaluating this expectation gives the following expression for the volatility futures price

$$F(V, T) = \frac{\alpha}{\beta}(1 - e^{-\beta T}) + Ve^{-\beta T}$$ \hfill (32)

The model represents futures prices as exponentially weighted averages of the current value of $V$ and the long run mean $\alpha/\beta$ of the risk adjusted process. As $T \to 0$, the futures prices converges to the current value of $V$. As $T \to \infty$, the futures price converges to $\alpha/\beta$. Since futures price is the expected value of $V$ taken with respect to the risk-adjusted process for $V$, the futures price will generally be a biased estimate of the actual expected future spot value of $V$.

Since $F(V, T)$ is not lognormally distributed in the Grünbichler and Longstaff model, the Black (1976) framework is not applicable in pricing VIX options. However, they provide a closed-form expression to price these options.

Let $C(V, K, T)$ denote the current value of a call option on $V$, where $K$ is the strike price and $T$ is the time until expiration. The call option can be expressed as

$$C(V, K, T) = \mathbb{E}[\max(0, V_T - K)]$$ \hfill (33)

Evaluating this expectation gives the following closed-form expression for the value of a volatility call
\[ C(V, K, T) = D(T)e^{-\beta T}VQ(\gamma K|\nu + 4, \lambda) \]
\[ + D(T)(\alpha/\beta)(1 - e^{-\beta T})Q(\gamma K|\nu + 2, \lambda) \]
\[ - D(T)Q(\gamma K|\nu, \lambda) \]  

(34)

Where \( Q(\cdot|\nu + i, \lambda) \) is the complementary distribution function for the non-central chi-squared density with \( \nu + i \) degrees of freedom and non-centrality parameter \( \lambda \). The volatility call price is an explicit function of \( V \) and \( T \), and depends on the exercise price \( K \), the riskless interest rate \( r \) through \( D(T) \), and parameters of the risk-adjusted volatility process \( \alpha \), \( \beta \) and \( \sigma \).

The major difference in the volatility call options valued with the Gr"{u}nbiicher and Longstaff model compared to call options on traded assets is that \( C(V, K, T) \) does not converge to 0 as \( V \) tends toward 0. This is related to the mean-reverting behavior of \( V \). If the value of \( V \) ever reaches zero, the process should immediately return to non-zero values. Consequently the value of the volatility call should be greater than zero when \( V = 0 \) since the call could still be in the money at expiration. As for calls on traded assets, if the traded asset equals zero, the probability that the price of the asset will eventually become greater than zero is null. This explains the third term in the volatility call equation instead of having only two terms like the Balck-Scholes (1973) equation. The extra term reflects that the call still have value even when \( V \) reaches zero.

Another interesting property is that the value of a volatility call becomes less than its intrinsic value for some value of \( V \). The price of the volatility call can be less than its intrinsic value when the call is only slightly in-the-money. This is again due to the mean reversion of volatility. When \( V \) is above its long-run mean, mean reversion implies that the expected future value of the volatility will be lower than its current value.

As with volatility futures prices, the volatility call price depend on \( V \) solely through the term \( e^{-\beta T} \). This implies that \( V \) will have little influence on the pricing of the call option as \( T \) increases. In this case, the delta of the call tends towards zero and the call value will be flat for a relevant range of \( V \). For a large enough \( T \), the deltas of all calls will approach zero.

Direct implication of these results suggests that longer-maturity calls should have little to no value as hedging instruments since their prices are not affected by changes in \( V \). However, even if longer-maturity calls have deltas near zero, shorter maturity calls can be used to hedge. There is nonetheless an upper-bound on the delta induced by the mean reversion effect, which could represent a significant restriction on the hedging properties of the short-term options.

In contrast to the Black-Scholes formula the volatility call option is not always an increasing function of \( T \). As \( T \) converges to infinity, the price of the call should be zero. Since \( V \) has a long-run stationary distribution, as \( T \) increases the value of \( D(T) \) used to discount the expected payoff approaches zero.
The effect of an increase in the strike price of a volatility call is always negative. An increase in $K$ should not have a symmetric effect to a decrease in $V$. An increase in $K$ has a significance on the prices of the long and short-term calls. A decrease in $V$ however, has little effect on the value of a long-term call.

The price of the volatility put $P(V, K, T)$ can be obtained with the following put-call parity,

$$P(V, K, T) = C(V, K, T) - D(T)F(V, T) + D(T)K$$ (35)

There is a difference between the put-call parity on volatility given with this model and the one on traded assets. The present value of a portfolio that pays $V$ at time $T$ is not equal to the current value of $V$. Instead it equals the present value of the futures price.

Volatility puts from the Gr"{u}nbichler and Longstaff model have similar properties to the volatility calls. The value of the put option can also be less than its intrinsic value. The delta of the put is a decreasing function of $T$. The relation between the put deltas and $T$ and the put deltas and $V$ mirrors the pattern for volatility calls. This again implies that longer maturity puts should be of limited use to the hedgers since delta approaches zero as $T$ increases.

4.2.2 Zhang & Zhu (2006)

In 2006, Zhang & Zhu produced their work in the more classic Heston (1993) framework. They investigated the VIX futures pricing using the Heston model (1993) where the SPX $S_t$ is modeled by following a diffusion process with stochastic instantaneous variance $V_t$.

$$dS_t = \mu S_t dt + \sqrt{V_t} S_t dW^S_t$$ (36)
$$dV_t = \kappa (\theta - V_t) dt + \sigma_v \sqrt{V_t} dW^V_t$$ (37)

Where $\mu$ is the expected return the SPX, $\theta$ is the long run mean level of the instantaneous variance, $\kappa$ is the mean reverting speed of the variance, $\sigma_v$ is the variance of the variance and $dW^S_t$ along with $dW^V_t$ are two standard correlated Brownian motion with constant correlation coefficient $\rho$. When changing probability measure $P$ to the risk neutral probability $Q$, the risk neutral measure for the SPX in then given by

$$dS_t = r S_t dt + \sqrt{V_t} S_t d\tilde{W}^S_t$$
$$dV_t = \kappa \theta - (\kappa - \lambda) - V_t) dt + \sigma_v \sqrt{V_t} d\tilde{W}^V_t$$ (39)
Where $d\tilde{W}_t^S$ and $d\tilde{W}_t^V$ are two new correlated standard Brownian motion with correlation coefficient $\rho$ and the parameter $\lambda$ is the risk premium associated to volatility.

With the VIX being defined as variance swap rate, they evaluate its value by computing the conditional expectation in the risk neutral measure.

$$VIX_t^2 = \mathbb{E}_t^Q \left[ \frac{1}{\tau_0} \int_t^{t+\tau_0} V_s \, ds \right]$$  \hspace{1cm} (40)

Where $\tau_0$ is 30 calendar days. Equation (39) yields

$$\mathbb{E}_t^Q[V_s] = \frac{\kappa \theta}{\kappa + \lambda} + \left( V_t - \frac{\kappa \theta}{\kappa + \lambda} \right) e^{-(\kappa + \lambda)(s-t)}$$  \hspace{1cm} (41)

Substituting this equation into equation (40) gives the following results and assuming the $VIX^2$ is linear function of instantaneous variance

$$VIX_t^2 = A + BV_t$$  \hspace{1cm} (42)

Where $A$ and B are functions of structural parameter, given by

$$A = \frac{\kappa \theta}{\kappa + \theta} \left[ 1 - \frac{1 - e^{-(\kappa + \lambda)\tau_0} \gamma}{(\kappa + \lambda)\tau_0} \right], B = \frac{1 - e^{-(\kappa + \lambda)\tau_0}}{(\kappa + \lambda)\tau_0}$$  \hspace{1cm} (43)

with $\tau_0 = 30/365$

Because the process of the VIX is observable, this can be used to back out the process of instantaneous variance.

The VIX futures can then be priced from the risk-neutral measure, the square root process of instantaneous variance in equation (39) determines the transition probability density

$$f^q(V_T|V_t) = ce^{-u-v} \left( \frac{u}{v} \right)^{q/2} I_q(2\sqrt{uv})$$  \hspace{1cm} (44)

where

$$c = \frac{2 + (\kappa + \lambda)}{\sigma^2 V(1 - e^{-(\kappa + \lambda)(T-t)})}, \ u = cV_t e^{-(\kappa + \lambda)(T-t)}, \ v = cV_T, \ q = \frac{2\kappa \theta}{\sigma^2 V} - 1$$

and $I_q(\cdot)$ is the Bessel function of the first kind of order $q$. The distribution function is the non-central chi-square, $\chi^2(2v; 2q + 2, 2u)$, with $2q + 2$ degrees of freedom and parameter of noncentrality $2u$ proportional to the current variance, $V_t$.

From this the VIX futures with maturity is given by

$$F_t = \mathbb{E}_t^Q(VIX_T) = \mathbb{E}_t^Q(\sqrt{A + BV_T}) = \int_0^\infty \sqrt{A + BV_T} f^Q(V_T|V_t) \, dV_T$$  \hspace{1cm} (45)
By estimating the parameters $\kappa$, $\theta$, $\sigma_v$, and $\lambda$ are then estimated by maximum likelihood function using the VIX historical data.

They tested the model by calculating the price of four VIX futures from the VIX spot and by calibrating the model with historical data using the maximum likelihood method. Results show that the model with the parameters estimated from the whole period (January 1990 to March 2005) overpriced the futures contracts by 16% for short term futures and 44% for the long term. By using one year of historical data, they however reduced discrepancy from 16% to 12% for the short-term futures and from 44% to 2% for the long-term.

4.2.3 Sepp (2008)

Building on a more sophisticated version of the Heston model (1993), Sepp (2008) addresses the problem of VIX futures and options pricing as trying to find a pricing method that would allow to fit the skew found in the VIX options. He uses a unified process for the SPX and VIX and the SPX realized variance with assumptions that the variance of returns of the SPX followed a square root diffusion process with jump diffusion. The introduction of the jump dynamics of the volatility implies the VIX options skew by assigning higher probabilities to larger values of the VIX in a short term scale which in turn verifies empirical results where negative SPX returns cause increased volatility as well as increased volatility of volatility.

To solve the pricing problem for a variety of volatility products, the joint dynamics of the asset price $S_t$, its variance $V_t$, and the realized variance $I_t$ are considered. The processes are given by

\begin{align*}
    dS_t &= rS_t dt + \sigma_t \sqrt{V_t} dW_t^S \\
    dV_t &= \kappa(1 - V_t) dt + \epsilon \sqrt{V_t} dW_t^V + J dN_t \\
    dI_t &= \sigma^2 V_t dt
\end{align*}

Where $\sigma_t$ is the deterministic ATM volatility, $\kappa$ is the mean reverting rate, $\epsilon$ is the volatility of volatility $dW_t^V$ along with $dW_t^V$ are two standard correlated Brownian motion with constant correlation coefficient $\rho$, $N_t$ is Poisson process with intensity $\gamma$ and $J$ is an exponentially distributed random jump with mean size $\eta$ and probability density function $\varpi(J) = \frac{1}{\eta} e^{-\frac{1}{2}\eta J}$.

The variance process $V_t$ is scaled to unity and $\sigma_t$ is assumed to be a piece-wise constant with local parameters chosen to match the term structure of the VIX futures. Using the generalized Fourier transforms, he obtains a closed-form solution for values of VIX options and futures.

The author tests his model by comparing hedging strategies between a variance delta-hedged portfolio and a variance delta-jump-hedged portfolio. The portfolio consists of a short VIX call and the calibrated model is used to provide hedging allocation in the corresponding VIX futures.
The results show that if the writer of a VIX call option follows the plain vanilla variance delta-hedging strategy, the portfolio remains short gamma in which small gains are offset by infrequent but rather huge losses when the VIX jumps. However, if the hedger follows the delta-jump-hedging strategy, the portfolio is practically gamma neutral. As a result, the P&L distribution under the last strategy is peaked at zero with small variations.

4.2.4 Lin & Chang (2009)

Similarly to Sepp(2008), Lin and Chang (2009) derived a closed-form indirect VIX option pricing model with jumps and an embedded stochastic volatility for the S&P500 index. They use four separate stochastic volatility processes with three of them as nested models on the main Eraker (2004) process. The main difference between Sepp(2008) and Lin & Chang’s model with the Eraker process is that the latter includes jumps in both the S&P500 and the S&P500 stochastic volatility.

They model the forward price of the S&P500, \( F_t(T) \), under the risk-neutral measure with a jump-diffusion process with stochastic instantaneous variance \( \nu_t \) given by

\[
\frac{dF_t(T)}{F_t(T)} = \sqrt{\nu_t}dW_{S,t} + J_t dN_t - \lambda_t \kappa dt
\]

\[
d\nu_t = \kappa_t (\theta - \nu_t) dt + \sigma_t \sqrt{\nu_t}dW_{\nu,t} + z_t dN_t
\]

where \( J_t = \exp(z_s) - 1 \) is the percentage price jump size with mean \( \kappa \). Satisfying the no-arbitrage condition, \( \kappa = \exp(\mu_j + \sigma_j^2/2)/(1 - \rho_j \mu_j) - 1 \). \( W_{S,t} \) and \( W_{\nu,t} \) are correlated Brownian motions with \( \rho dt = \text{corr}(dW_{S,t}, dW_{\nu,t}) \) and are independent of the compound Poisson processes \( z_s dN_t \) and \( z_\nu dN_t \). The instantaneous variance \( \nu \) follows a mean reverting square-root process with exponentially distributed jump size \( z_\nu \) that is correlated with price jump size \( z_S \) through \( z_S = \mu_j + \rho_j z_\nu \). Jumps in volatility are assumed to have an exponential distribution, \( z_\nu \sim \exp(\mu_\nu) \) whereas jumps in asset prices are normally distributed conditional on the realization of \( z_\nu \).

The mean of \( z_S \) is \( \mathbb{E}(z_S) = \mu_j + \rho_j \mu_\nu \), with variance \( \text{var}(z_S) = \sigma_j^2 + \rho_j^2 \mu_j^2 \) and is correlated with \( z_\nu \) through \( \rho_j \mu_\nu / \sqrt{\sigma_j^2 + \rho_j^2 \mu_j^2} \). The underlying return and its volatility share the same jump arrival uncertainty followed by a Poisson process \( N_t \) with state-dependent intensity \( \lambda_t = \lambda_0 + \lambda_1 \nu_t \). The speed adjustment is given by \( \kappa_\nu - \lambda_1 \mu_\nu \), the long run mean is \( (\kappa_\nu \theta_\nu + \lambda_0 \mu_\nu) / (\kappa_\nu - \lambda_1 \mu_\nu) \) and the variation of the instantaneous variance \( \nu \) is \( \sigma_\nu^2 \nu_t + 2\lambda_1 \mu_t^2 \).

With the square of VIX being defined as the variance swap rate, it can be evaluated by computing the conditional expectation under the risk-neutral measure \( Q \) with

\[
VIX_t^2 = \frac{1}{\tau} \mathbb{E}_t^Q \left[ \int_t^{t+\tau} (\nu_u du + J_u^2 dN_u) \right]
\]
where τ is 30 calendar days by definition.

The model can be then calibrated by fitting the parameters to obtain the day’s option prices.

Results obtained with this model are however inconclusive as to the internal consistency of the implicit parameter by comparing the model’s implied parameters with those implicit in the time series of the VIX futures prices. The main culprit causing the significant misspecification of the model is the implausible levels of implied volatility variation of forward VIX to rationalize observed option prices. However, on a relative scale it is somewhat less misspecified when comparing to the Heston model. The results are also less than convincing when looking at out-of-sample option pricing. The test is done with the previous month structural parameters and the current day’s VIX and VIX futures prices to calculate the current day’s VIX option prices. The mean absolute pricing error (MAE) is employed as evaluation metric. With pricing performance significantly increased in comparison to the Heston model, the pricing error is still significant.

It is however important to mention that in comparison to other equilibrium models, the Lin & Chang’s model with the Eraker process has best performance, which shows that modelling volatility jumps should be important.

4.3 No Arbitrage Models

The No-Arbitrage framework is commonly used in the pricing of complex interest rates derivatives. As mentioned earlier, the approach does not try to explain term-structure but uses it to price higher order derivatives. In 1990 David Heath, Bob Jarrow, and Andy Morton (HJM) published an important paper describing the No-Arbitrage condition that must be satisfied by a yield curve (Heath et al., 1992). Considering the similarities between forward interest rates and forward variance, Dupire (1992) adapted their work to the variance term-structure. The approach basically consists in taking today’s variance curve and adding a stochastic process where the drift should follow exactly today’s given term-structure. Buehler (2006) later presented the adaptation of the framework to variance curve in a more general framework.

4.3.1 Dupire (1993)

Dupire (1993) uses the HJM framework in order find a way to price claims contingent on either the Spot price, its volatility or both. From the given prices on call options, he derives an arbitrage free value of forward variance which in turns allows pricing more exotic contingent claims.

Assuming zero interest-rate to ease the presentation and a continuum of call options being traded which the prices at time 0 are consistent with no arbitrage,
then for all $T$ the knowledge of call prices across all moneyness allows to obtain $\phi_T$, the implicit risk-neutral distribution of $S_T$ which in turn determines the price of the European call options.

Considering a portfolio of calls (known as a butterfly) that includes

$$\frac{C_{K-\epsilon,T} - 2C_{K,T} + C_{K+\epsilon,T}}{\epsilon^2}$$

(49)

As $\epsilon$ tends toward zero, its profile converges to a Dirac function at point $K$, and its price converges to $\phi_T(K)$ if finite. From this, $\phi_T$ can be expressed as

$$\phi_T(K) = \frac{\delta^2 C_{K,T}}{\delta K^2}$$

(50)

Denoting $P_T$ the associated probability which density function is $\phi_T$, a price at time 0 of any payoff profile $f$ delivering $f(S_T)$ at maturity $T$ in $L^1(P_T)$ can be given by

$$f_T(0) = \int_0^\infty f(S_T)\phi_T(S_T)dS_t \equiv \mathbb{E}^{P_T}[f]$$

(51)

Particularly, taking a claim $L_T$ delivering the logarithm of $S_T$ at time $T$ that belongs to $L^1(P_T)$, then

$$L_T(0) = \mathbb{E}^{P_T}[\ln(S_T)]$$

(52)

Which under appropriate integrability conditions is the only information needed to obtain the arbitrage free value of the forward variance.

Introducing filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_T, P)$, where $\mathcal{F}_T$ is a right continuous filtration containing all $P$-null sets, $\mathcal{F}_0$ being trivial, and a Spot price process $S_t$ driven by the following stochastic differential equation:

$$dS_t/S_t = \mu_t dt + \sigma_t dW_{1,t}$$

(53)

Where $W_1$ is a $P$-Brownian motion adapted to $\mathcal{F}_t$, and $\mu_t$ and $\sigma_t$ may in turn be stochastic processes measurable and adapted to $\mathcal{F}_t$ then by Itô’s Lemma

$$d\ln S_t = \frac{dS_t}{S_t} - \frac{\sigma_t^2}{2}dt$$

(54)

Integrating between $T_1$ and $T_2$

$$\ln S_{T_2} - \ln S_{T_1} = \int_{T_1}^{T_2} \frac{dS_t}{S_t} - \frac{1}{2} \int_{T_1}^{T_2} \sigma_t^2 dt$$

(55)

Which can be rewritten in terms of $\sigma_t^2$

$$\int_{T_1}^{T_2} \sigma_t^2 dt = 2\int_{T_1}^{T_2} \frac{dS_t}{S_t} - 2(\ln S_{T_2} - \ln S_{T_1})$$

(56)
The stochastic integral \( \int_{T_1}^{T_2} \frac{dS_t}{S_t} \) can be interpreted as the wealth at time \( T_2 \) of a strategy consisting of permanently keeping one unit of a riskless bond \( B \) invested in the risky asset \( S \) between time \( T_1 \) and time \( T_2 \) through the possession \( \frac{1}{S_t} \) units of \( S \). This strategy is self financing since the interest rate is zero.

The left hand side of equation (56) is the cumulative instantaneous variance of the Spot return between \( T_1 \) and \( T_2 \) and can be reproduced by the portfolio \( 2(L_{T_2} - L_{T_1}) \) associated with a dynamic self-financing strategy consisting in keeping 2 units of riskless asset \( B \) permanently invested in the risky asset \( S \) between \( T_1 \) and \( T_2 \).

Considering a variance swap contract delivering \( \int_{T_1}^{T_2} \sigma^2_t \, dt \) at \( T_2 \) traded at time \( t \), by arbitrage it should have a unique possible value equal to \( 2(L_{T_1}(t) - L_{T_2}(t)) \). In the event that there is no such forward market, this contract can be synthesised which validates the assumption of its existence.

Let \( V_T \) be a forward variance swap contract on the instantaneous variance to be observed at time \( T \). From the previous development, the value of the contract \( V_T \) at any time \( t < T \) can be obtained by differentiation where

\[
V_T(t) = -2 \frac{\partial L_T(t)}{\partial T} \tag{57}
\]

As \( -L_T \) is a convex function of \( S \), \( V_T(t) \) is positive. From the values of \( (L_T(0))_T \), the initial instantaneous forward variance curve is given by,

\[
V_T(0) = -2 \frac{\partial L_T(0)}{\partial T} \tag{58}
\]

From the results above and following the lines of HJM to directly model the forward variance, this should automatically ensure compatibility \( (L_T(0))_T \). Assuming that \( V_T(t) \) is Lognormal leading to

\[
\frac{dV_T(t)}{V_T(t)} = adt + bdW_{2,t} \tag{59}
\]

where \( a \) and \( b \) are constant and \( W_2 \) is a second Brownian motion adapted to \((\mathcal{F}_t)_t\) and possibly correlated with \( W_1 \). The risk-neutral process for the forward variance is then given by defining \( dW_{2,t}' = dW_{2,t} + \frac{a}{b} \, dt \) which enables to rewrite equation (59) as

\[
\frac{dV_T(t)}{V_T(t)} = bdW_{2,t}' \tag{60}
\]

where \( W_2' \) is a Brownian motion under \( Q_2 \), the \( P \)-equivalent probability classically obtained by Girsanov’s Theorem. Applying Itô’s lemma to equation (60) yields,

\[
d\ln V_T(t) = -\frac{b^2}{2} \, dt + bdW_{2,t}' \tag{61}
\]
Integrating between 0 and \( t \) leads to

\[
d \ln V_T(t) = \ln V_T(0) - \frac{b^2}{2} dt + bdW_{2,t}'
\]  
(62)

Under \( Q_2 \), \( V_T \) is a martingale, which means

\[
\mathbb{E}^{Q_2}[V_T(t)|V_T(0)] = V_T(0)
\]  
(63)

The no-arbitrage process for the instantaneous variance at time \( t \), \( v_t \equiv V_t(t) = \sigma_t^2 \), is then given by,

\[
d \ln v_t = \left( \frac{\partial \ln V_T(0)}{\partial t} - \frac{b^2}{2} \right) dt + bdW_{2,t}'
\]  
(64)

and for instantaneous volatility,

\[
d \ln \sigma_t = \frac{1}{2} \left( \frac{\partial \ln V_T(0)}{\partial t} - \frac{b^2}{2} \right) dt + \frac{b}{2} dW_{2,t}'
\]  
(65)

A one factor model that is free of arbitrage is obtained from the previous development. This means however that all forward variances are perfectly correlated and any one of the them could theoretically be used to hedge volatility risk. This also implies that the volatility curve only moves in parallel shifts. For those reasons it becomes obvious that a one factor assumption is unrealistic for practitioners since the volatility curve observes a wide variety of twists.

Even though the use of one factor seems overly simplistic, Dupire (1993) was however a stepping stone in arbitrage-free variance curve modelling.

### 4.3.2 Buhler (2006)

Buhler presented in 2006 a more general approach to modelling variance curves within the HJM framework that can include more than one factor.

Within probability space \((\Omega, \mathcal{F}, \mathcal{F}_T, P)\) with a \(d\)-dimensional Brownian motion \( W = (W_j)_{1,\ldots,d} \), being consistent with previous development and defining the variance swap \( V_t(T) \) by,

\[
V_t(T) = \mathbb{E}^{P_T} \left[ \int_0^T v_s ds \right]
\]

Given that \( V(T) \) is a martingale, a measure \( \varphi(T) \) can be found such that

\[
V_t(T) = V_0(T) + \sum_{j=1}^d \int_0^t \varphi_j^2(T) dW_s^j
\]  
(66)
Taking the derivative along $T$ gives the fixed maturity $T$-forward variance seen at time $t$

$$v_t(T) = \partial_T V_t(T) = v_0(T) + \sum_{j=1}^{d} \int_0^t \beta_j(T) dW_j^j$$

(67)

Where $\beta(T) = \partial_T \varphi(T)$.

In order to have a variance curve model that is free of arbitrage, the forward variance has to remain positive. A way to ensure the respect of this constraint is to make the assumption that the variance is log-normally distributed and the process for $v_t(T)$ is given by,

$$dv_t = -\frac{1}{2} \sum_{j=1}^{d} \beta_j^2(T) dt + \sum_{j=1}^{d} \beta_j^j(T) dW_j^j$$

(68)

This is very similar to the forward variance process given by Dupire (1993) with the exception that there is now more than one Brownian Motion driving the forward variance curve. Extra care needs to be taken when choosing the parameters $\beta_j^j$ in order to ensure that $v_t(T)$ and $V_t(T)$ remain martingales.

4.3.3 Huskaj & Nossman (2013)

Huskaj and Nossman (2013) derived a one factor model for the VIX term structure that is consistent with key-results from the HJM framework where the risk neutral-drift is a function of volatility and in which the VIX futures price dynamics are specified exogenously. They use a multi factorial model capturing empirical properties of VIX futures returns by assuming that they are normal inverse Gaussian (NIG).

When looking at empirical VIX futures return, the distribution shows statistically significant positive skewness as well as excess kurtosis across all maturity. Hence, for a more flexible distribution that fits empirical returns properties, they choose to model the VIX futures returns with a NIG distribution rather than a normal distribution.

The model starts from defining the VIX futures log price under the physical probability measure as

$$d \ln(F_t(T)) = a_t(T) dt + \sigma_t(T) dN_t(T)$$

(69)

where $F_t(T)$ is the VIX futures price at time $t$ with maturity $T$, $a_t(T)$ and $\sigma_t(T)$ are the instantaneous drift and volatility functions of the VIX futures logarithmic price, and $N_t(T)$ is the stochastic process under the NIG distribution with density function given by

$$f(x; \alpha, \beta, \mu, \delta) = \frac{\delta \alpha}{\pi} \exp(\delta \sqrt{\alpha^2 - \beta^2} + \beta(x - \mu))$$

$$\times \frac{K_1(\alpha(\sqrt{\delta^2 + (x - \mu)^2}))}{\sqrt{\delta^2 + (x - \mu)^2}}$$

41
where $0 \leq |\beta| < \alpha$, $\mu \in \mathbb{R}$ is the location parameter, $\delta > 0$ is the scale parameter, and $K_1(\cdot)$ is the modified Bessel function of third order and index one, that is, $K_1(x) = \int_0^\infty \exp(-xcosh(t))cosh(t)dt$. The parameters $\alpha$ and $\beta \in \mathbb{R}$ give the distribution’s kurtosis and skewness, respectively where a smaller value of $\alpha$ implies fatter tails and $\beta = 0$ implies symmetry. The NIG distribution converges to the normal when $\beta = 0$, $\alpha \to \infty$, and $\delta/\alpha = \sigma^2$. $N_t(T)$ is a NIG distributed process if $N_1(T)$ is NIG distributed. The mean variance, skewness and kurtosis of the NIG distributed $N_1(T)$ under the physical measure are given by

$$E(N_1(T)) = \mu + \delta \beta \sqrt{\alpha^2 - \beta^2} \quad (70)$$

$$V(N_1(T)) = \frac{\delta \alpha^2}{(\alpha^2 - \beta^2)^{3/2}} \quad (71)$$

$$Skew(N_1(T)) = \frac{3\beta}{\alpha \sqrt{\delta \sqrt{\alpha^2 - \beta^2}}} \quad (72)$$

$$Kurt(N_1(T)) = 3 + \frac{3(1 + 4(\beta/\alpha)^2)}{\delta \sqrt{\alpha^2 - \beta^2}} \quad (73)$$

The moments are time invariant. To define $a_t(T)$ and $\sigma_t(T)$ consistently, the process $N_t(T)$ has to have zero mean and a unit variance. This can be achieved by setting the location and scale parameters $\mu$ and $\delta$ accordingly in equations (71) and (72).

The volatility function $\sigma_t(T)$ in equation (69) is specified as

$$\sigma_t(T) = \sigma_1 \exp(-\sigma_2(T-t)) + \sigma_3 \quad (74)$$

where $\sigma_1$, $\sigma_2$, and $\sigma_3$ are all non-negative. The argument for the choice of volatility function specification in equation (74) is that it seems to be a good candidate to capture the empirical volatility of volatility term structure since it rises sharply towards $\sigma_1 + \sigma_3$ as the contract approaches maturity and decays toward $\sigma_3$ as the maturity increases.

For the VIX futures price to remain a martingale under the risk-neutral measure $Q$, adjustments have to be made for the drift term $a_t(T)$ from equation (69) which can be defined as

$$\exp\left(-\int_t^T a_s(T)ds\right) = \mathbb{E}^Q\left[\exp\left(\int_t^T \sigma_s(T)dN_s(T)\right)\right] \quad (75)$$

$$= \exp\left(\int_t^T \psi_s(-i\sigma_s(T),T)ds\right)$$

For VIX options pricing, the model presented so far is incomplete since the driving noise for the VIX futures dynamics is a NIG process. The risk neutral probability measure derived from the Esscher transform, which enables NIG
and Weiner processes to remain in the same families when changing measures. This in turn makes the market price of risk a necessary input. The only change when going from the original NIG process to its Esscher transform is that $\beta$ becomes $\beta + \theta_1(T)$ where $\theta_1(T)$ parametrizes the time-dependant risk premium. Thus under the risk-neutral Esscher measure, the Esscher transform of the NIG process stays a NIG process with $\beta + \theta_1(T)$ specified as 

$$\tilde{\beta}_t(T) = \beta + \theta_1(T) = \beta - \theta_1 + 1/\theta_2 + T - t$$

(76)

and the characteristic exponent specified as 

$$\psi_t(u, T) = iu\mu + \delta \left( \sqrt{\alpha^2 - \tilde{\beta}_t^2(T)} - \sqrt{\alpha^2 - (\tilde{\beta}_t(T) + iu)^2} \right)$$

(77)

From empirical data, the model is then calibrated using the maximum likelihood procedure to find parameters $\alpha$, $\beta$, $\sigma_1$, $\sigma_2$, $\sigma_3$, $\theta_1$, and $\theta_2$. They then test their model for VaR estimates which are evaluated using the failure rate $f$ and the Kupiec (1995) LR test. The model with the NIG distribution seems to perform well in and out-of-sample when compared to models using a normal distribution for VIX futures return. When analysing the in-sample tail-risk of a long and a short VIX futures position, the NIG model seems to be more adequate according to the LR test at any VaR level ($\alpha$) for both positions while normal models severely overestimate the risk at $\alpha = 5\%$ for a long position and underestimate the risk at $\alpha = 2.5\%$ and $\alpha = 1\%$ in a short position. The out-of-sample results show that the NIG and normal models performs well for a long VIX futures position, except at $\alpha = 5\%$, where the null hypothesis of the LR test is rejected for the at the 1% level. As for a short position, the null is not rejected for any of the NIG models at the $\alpha = 5\%$ level, whereas all normal models again gravely underestimate the risk exposure at equal to $\alpha = 2.5\%$ and $\alpha = 1\%$.

### 4.3.4 Cont & Kokholm (2013)

Cont and Kokholm (2013) propose and study a model that includes joint dynamics of an index and a set of forward variance swaps rates written on the index that allows the pricing of volatility derivatives and options on the underlying index to be consistent. Similar to Buehler (2006), the building bloc of the model is the forward variance process in which they have added a stochastic jump in both the variance swap dynamics and index dynamics. This in turn allows the model to fit simultaneously the volatility skews in the index and volatility options while at the same time reproducing a realistic representations of empirically observed properties of variance swap dynamics by allowing jumps in volatility and returns.

The basics of their model for variance swap dynamics is similar to equation 68 with one Brownian motion and to which is added a stochastic jump which yields the following process for forward variance $v_t(T)$ from $T_n$ to $T_{n+1}$.
\[ v_t(T) = v_0(T) e^{X_t(T_i)} \tag{78} \]
\[ = v_0(T) \exp \left[ \int_0^t \mu_s(T_i) ds + \int_0^t \omega e^{-k_1(T_i-s)} dZ_s + \int_0^t \int_{\mathbb{R}} e^{-k_2(T_i-s)} x J(dx ds) \right] \]

where \( J(dx dt) \) is a Poisson random measure with compensator \( \nu(dx) dt \) and \( Z \) is a Weiner process independent of \( J \). For \( v_t(T) \) to be a martingale, this imposes the following condition,
\[ \mu_\ell(T_i) = -\frac{1}{2} \omega^2 e^{-2k_1(T_i-t)} - \int_{\mathbb{R}} \left( \exp \left[ e^{-k_2(T_i-t)} x \right] - 1 \right) \nu(dx) \tag{79} \]

For \( t > T_i \), \( v_t(T_i) = v_{T_i}(T_i) \) and where \( \omega \), \( k_1 \), and \( k_2 \) are constants that are calibrated to fit the model to market options prices.

This specification allows for a jumps and the exponentials inside the integral allow to control the term structure of volatility of volatility.\(^8\) For example if \( k_1 \) is large and \( k_2 \) is small, the diffusion term \( Z \) will affect the short term end of the curve, while the jumps impact the entire curve. The exponential functions for the volatility and jump specifications may be changed for added flexibility but are important in order to use an affine specifications that allow Fourier-based pricing of European volatility derivatives.

The same specifications can be used for VIX futures, hence the model for the forward variance and VIX futures may both be calibrated with VIX options. There is however a slight adjustment for convexity to be made. This gives
\[ VIX_t(T_i) = \sqrt{v_t(T_i)} = \sqrt{v_0(T_i) M_t(T_i)} \tag{80} \]

Where \( M_t(T_i) \) is an exponential martingale given by
\[ M_t(T_i) = \exp \left[ \int_0^t \eta_s(T_i) ds + \frac{1}{2} \int_0^t \omega e^{-k_1(T_i-s)} dZ_s + \int_0^t \int_{\mathbb{R}} \frac{1}{2} e^{-k_2(T_i-s)} x J(dx ds) \right] \tag{81} \]

where
\[ \eta_\ell(T_i) = -\frac{1}{8} \omega^2 e^{-2k_1(T_i-t)} - \int_{\mathbb{R}} \left( \frac{1}{2} \exp \left[ e^{-k_2(T_i-t)} x \right] - 1 \right) \nu(dx) \tag{82} \]

Having a model for the VIX futures from which they can back out the process for the forward variance, they can then calibrate their model to price consistently for the index options.

The results that they obtain show that the model is able to fit option prices across time to maturity and moneyness (S&P500 and VIX) with an average error that is less than 1% on the market price.

\(^8\)Note that the volatility specification is similar to Huskaj and Nossman (2013) from equation 74 but leaving out the \( \sigma_3 \) term for the long term average.
5 Model Specifications

The previous section presented various VIX models. Since most of the previous studies on VIX option pricing models did not have any options prices available at the time, Wang and Daigler (2011) tested and compared those models. They found that the Lin & Chang (LC) model, which is the best performing model within the Heston framework, does not offer comparable performance to either the Whaley, Carr & Lee (CL) or Grünbichler & Longstaff (GL) models.

They tested the models by looking at the relative and cash error (signed and unsigned) across time to expiry and moneyness during the period of January through September 2007. They have found that the Whaley model and the CL model had similar performance to the GL model for almost all in-the-money options across all expiries. However, only the Whaley model has similar performance to GL for near-the-money options. The GL model also showed smaller error relative to both Whaley and CL in terms of percentage error and in terms of cash error for the out-of-the-money options and deep-out-of-the-money options. They also found that the GL model outperforms the CL model across all time to expiration except for deep-in-the-money options when comparing the pricing error across time to expiration. The Whaley model however produces smaller error for longer maturities (>90 days) than the GL model. For shorter maturities, the Whaley model performs better for in-the-money options and the GL model performs better for out-of-the-money options.

The conclusion of Wang and Daigler (2011) suggests that the VIX generates different implied volatility surfaces than stock index options and therefore it is better to model the underlying process of the VIX directly rather than modelling the volatility of the SPX as it is done within the Heston framework. They also find that there is no direct VIX model that clearly outperforms the others in pricing options across all moneyness and maturities. Following those results, they suggest that simpler might be better when it comes to the choice of a VIX model.

Following Wang & Daigler’s opinion the choice of our VIX model should go towards simplicity. In order to make the right choice of model, let's define our needs. Seemingly, the shape and dynamics of the VIX futures curve play a great role in the VIX futures return. Keeping in mind that we are not interested in pricing the VIX futures but evaluating more complex products or risk metrics from the futures prices, hence the initial curve should be taken for granted. This means that a suitable model should fit the initial VIX curve perfectly. This can be easily achieved in the no-arbitrage framework. We also want to allow the curve to take different shapes which implies that there is a need for a multi-factorial model. Using a multi-factorial forward variance that is distributed log-normally, this will allow to back out a process for the VIX futures curve at...
any time $t$ and use Monte-Carlo simulation in order to solve for each maturity $T$. The approach that is used is derived from the general framework from Buehler (2006) and can also be seen as a variation from Cont and Kokholm (2013).

The variance or volatility swap curve models as Dupire (1992), Buehler (2006), and Cont and Kokholm (2013) take into account that the variance swap $V_t(T)$ is a martingale. This means that the expected value at time $t = T$ of the variance swap $E[V_t(T)] = V_0(T + t)$. This corresponds to the cost-of-carry model where $E[S_T] = F_0(T)$ or in the case of the VIX, $E[VX_{T}] = VX_0(T)$. Thus implying that at time $t = T$ the spot VIX is expected be worth its future value $VX_0(T)$. This approach works for VIX options valuation since it is made within the risk-neutral framework where the futures contract on which the option is written is a martingale.

If we want to model the dynamics of the VIX futures, we have seen in the literature that the empirical properties demonstrate that instead the expected value $E[VX_{T}(T)] = VX_0$ since the futures roll down along the curve converging to the spot VIX. Thus meaning that constant tenors of the VIX curve are martingales hence $E[VX_{t}(T)] = VX_0(T - t)$. This allows the VIX futures contract to converge towards the spot VIX as $t$ approaches $T$.

The argument as to why we should consider the constant tenors $VX_t(T + t)$ to be martingales instead of the futures contract themselves is the following.

Considering that the spot VIX as well as the VIX futures are portfolios of S&P500 options Vega exposure that is constant across moneyness. When referring to equation (9), this makes the portfolio a function of the square-root of time to maturity and volatility. As we have seen in previous sections, the spot VIX is constantly rebalanced between the first and second S&P500 options maturity in order to achieve a constant time to maturity of 30 days. Thus, with everything else being equal (i.e. no change in the S&P500 options implied volatility), even as time passes the expected value of the Vega portfolio will not change which yields that $E[VX_{t}(T - t)] = VX_0(T)$ where $T = 30$ days. The same reasoning can be applied with any constant maturity.

Keeping in mind that the constant maturity $VX_t(T)$ is a martingale, this implies that the constant maturity forward variance $v_t(T)$ is a martingale as well. As we have said earlier we want to model the forward variance and then integrate it in order to have our variance swap (VIX) price. The reasons as to whether we choose to model the VIX from the forward variance and not back out the forward variance from the VIX process as in Cont and Kokholm (2013) is in regards to using more than one factor to drive the curve dynamics. We use multiple factors since we want to allow the curve to move as freely as possible. This means that the different maturities are correlated by a factor $\rho$. However, allowing the VIX curve to have multiple shapes can also cause calendar arbitrage (negative forward variance). By modeling the forward variance dynamics and then integrating to obtain the VIX curve, we then insure that the forward variance remains positive.
The use of multiple factors also affects the volatility of volatility specifications. The use of only one factor as in Dupire (1992) and Cont and Kokholm (2013) implies that all maturities are perfectly correlated ($\rho = 1$), hence the forward variance volatility is the same as the variance swap volatility. This is not the case when the correlations between maturities are less than 1. This will be discussed further on in section 5.3.

5.1 Multi-Factorial No-Arbitrage Model for VIX Futures

Following Buhler’s (2006)\(^{10}\) approach for consistent variance curve within the HJM framework, we make the assumption that the variance swap $V_t(T)$ is log-normally distributed which leads to

$$\ln(V_t(T)) = \ln \left[ V_0(T) - \frac{1}{2} \sum_{j=1}^{d} \int_0^t \varphi_{s_1}^j(T)^2 ds_1 + \sum_{j=1}^{d} \int_0^t \varphi_{s_1}^j(T)dW_{s_1}^j \right]$$

(83)

Keeping in mind that $\partial_T V_t(T) = v_t(T)$, and $\partial_T \varphi_1^j(T) = \beta_t^j(T)$ we have

$$\ln(V_t(T)) = \ln \left[ \int_0^T v_t(s_2)ds_2 \right]$$

$$= \ln \left[ \int_0^T v_0(s_2)ds_2 - \frac{1}{2} \sum_{j=1}^{d} \int_0^t \beta_{s_1}^j(s_2)^2 ds_1 + \sum_{j=1}^{d} \int_0^t \beta_{s_1}^j(s_2)dW_{s_1}^j \right]$$

(84)

In our case $\beta_t^j(T)$ is a deterministic function (see following sections), knowing the VIX is the square root of a variance swap, i.e. $\ln[VIX_t(T)] = \ln[\frac{1}{2} V_t(T)]$, the VIX futures prices at time $t$ can be given by,

$$\ln[VIX_t(T)] = \ln \left[ VIX_0(T) - \frac{1}{8} \sum_{j=1}^{d} \int_0^t \varphi_{s_1}^j(T)^2 ds_1 + \int_0^T \sum_{j=1}^{d} \frac{\beta_{s_1}^j(s_2)}{2} dW_{s_1}^j \right]$$

(85)

And the process for $\ln[VIX_t(T)]$ given by,

$$d\ln[VIX_t(T)] = \ln \left[ -\frac{1}{8} \int_0^T \sum_{j=1}^{d} \beta_{s_1}^j(s_2)^2 dt ds_2 + \int_0^T \sum_{j=1}^{d} \frac{\beta_{s_1}^j(s_2)}{2} dW_{s_1}^j \right]$$

(86)

The next section defines $\beta_t^j(T)$ by using a principal component analysis.

\(^{10}\)The notation that is used from this point forward will be consistent with Buehler (2006).
5.2 Curve Dynamics

A good deal of authors in the interest rates literature, notably Litterman and Scheinkman (1991), Rebonato (1996), Scherer and Avellaneda (2002) and Driessen et al. (2003) use the principal component analysis method in order to model the term-structure dynamics within the Heath-Jarrow-Morton framework.

The main idea behind the method is to look at the curve as a whole instead of correlated individual maturities. This in turn brings dimensionality reduction of the problem.

As an example, the US term-structure has a set of ten maturities. When dealing with this problem, one would have to simulate ten series of correlated random numbers. In this context, using the PCA method to approximate the dynamics of the curve would lead to simulating only three series\footnote{When modeling interest rates term-structure a number of three factors is usually regarded as a good approximation.} of random numbers which reduces computing time significantly.

The similarities between forward variance and forward interest rates should enable to use this method to extract the main factors driving the evolution of the forward variance term structure. Following Lu and Zhu (2010) which studied the VIX futures curve using the PCA method, it seems that the first three factors are statistically significant in order to capture the dynamics of the curve correctly.

Using historical VIX futures data\footnote{Data obtained from Bloomberg LP.}, the principal component analysis is done on the correlation matrix of the log returns of the forward variance term structure.

Let $v_{t_k}(T_i)$ be the forward variance at maturity $T_i$ where $i = 0, ..., n$ and let $r_{t_k}(T_i)$ be the log return of the forward variance with maturity $(T_i)$ at time $t_k$ be given by

$$r_{t_k}(T_i) = \log\left(\frac{v_{t_k}(T_i)}{v_{t_k-1}(T_i)}\right)$$

(87)

Let $\Sigma$ be the covariance matrix of the log returns of the forward variance term structure where,

$$\Sigma_{i,j} = \mathbb{E}\left[\left(r_{t_k}(T_i) - \mathbb{E}[r_{t}(T_i)]\right)\left(r_{t_k}(T_j) - \mathbb{E}[r_{t}(T_j)]\right)\right]$$

(88)

For $i = 0, ..., n$ and $j = 0, ..., n$.

Taking the eigenvalues and eigenvectors for $\Sigma$ will allow to obtain factor loadings for each maturity.

Let $\Phi_i$ be the $i^{th}$ eigenvector corresponding to $\lambda_i$, the $i^{th}$ largest eigenvalue of $\Sigma$ and $\beta_i(T_j)$ be the $i^{th}$ component for maturity $T_j$, the three factor loadings
corresponding to each maturity are then given by

$$\beta_i(T_j) = \sigma_v(T_j) \sqrt{\lambda_i(\Phi_2^2)_j}$$  \hspace{1cm} (89)

With this result, the only remaining parameter left to define is $\sigma_v(T)$ which is presented in the next section.

5.3 Forward Variance Volatility

Using the HJM framework intuition, we want to minimize parameter estimations/calibrations and take as much information from the market as possible. Considering that there is an option market for the VIX that is fairly liquid, this offers information in regards to the VIX futures and the forward variance implied distribution. By taking the implied distribution from the market, we are taking the consensus from all the participants in the market in regards to the estimation of the implied volatility. This line of thought also goes along with the price discovery theory in the sense that the aggregated information from all the participants should be more accurate on the given price of an asset than each individual.

Furthermore, literature on volatility forecasting using implied volatility seems to go towards the same conclusions as to whether or not implied volatility produces quality forecast on the realized volatility of an index. Results from Fleming (1998) indicating the implied volatility on the S&P 100 to be an upward biased forecast that still contained relevant information regarding future volatility, but further studies such as Christensen and Prabhala (1998), Christensen and Strunk Hansen (2002), Corrado and Miller (2005) argue that the biases were in fact induced by econometric problems in variables and draw the conclusions that implied volatility and implied volatility indexes are in fact an efficient forecasts for the realized volatility on stock indexes.

There is however no known literature on the implied-realized volatility relation of the VIX, which we will leave to future research. We will thus assume the efficiency of the VIX options market from its abundant liquidity and make the hypothesis that the VIX implied volatility is also an efficient forecast of its realized volatility. This could be achieved by replicating the previous studies with the VVIX index which is the 30 day volatility swap on the VIX (VIX of VIX).

Following these principles and using the CBOE methodology, we first start by extracting the variance swap price for each maturity $T$ from the VIX options by using equation (2). This gives us the average implied variance across moneyness for the VIX futures for each expiry $T$. From this information, we can now find the forward variance of $v(T)$, $\sigma_v(T)$ that will allow to simulate the VIX futures curve with respect to its variance term-structure.
We first need to get the variance of the \( VIX^2 \) which is given by \( \sigma_{VIX^2}^2(T) = 4\sigma_{VIX}^2(T) \). We have seen from the previous sections that the variance swap is the integrated forward variance from time 0 to time \( T \) such that

\[
\mathbb{E}[V(T)] = \mathbb{E} \left[ \int_0^T v(s) \, ds \right] \tag{90}
\]

Hence,

\[
\mathbb{E}[\sigma_{VIX^2(T)}^2] = \mathbb{E} \left[ \int_0^T \sigma_v^2(s) \, ds \right] \tag{91}
\]

Solving for \( \sigma_{VIX^2(T)}^2 \) and discretizing the equation we obtain,

\[
\sigma_{VIX^2}^2(T_n) = \sum_{i=1}^n \left( \frac{T_i - T_{i-1}}{T_n} \right)^2 \sigma_v^2(T_{i-1,i}) \nonumber
\]

\[
\quad + 2 \sum_{i=1}^n \sum_{1 \leq j < k \leq n} \rho_{i,j} \frac{T_i - T_{i-1}}{T_n} \frac{T_j - T_{j-1}}{T_n} \sigma_v(T_{i-1,i}) \sigma_v(T_{j-1,j}) \tag{92}
\]

Having already defined \( \rho_{i,j} \) in the previous section, we solve by iteration to find \( \sigma_v(T_{i-1,i}) \) for each maturity.

We can now present the VIX model from equation (86),

\[
d\ln[VIX_t(T)] = \ln \left[ -\frac{1}{8} \int_0^T \sigma_v^2(s_2) \sum_{j=1}^d \lambda_j(\Phi_j^2)_{s_2} \, dt \, ds_2 \right.
\]

\[
\quad + \left. \int_0^T \frac{\sigma_v(s_2)}{2} \sum_{j=1}^d \sqrt{\lambda_j(\Phi_j^2)_{s_2}} \, dW_{j,s_2} \right] \tag{93}
\]

We now have a VIX curve model that allows different shapes for the term-structure that also fits the initial curve perfectly. We can also use the model in a joint process with the S&P500 index in order to obtain a complete market including the S&P500 and the VIX futures. This can be achieved by choosing a correlation structure between \( S_t(T) \) and \( V_t(T) \) that will allow \( S_t(T) \) to be a martingale and thus fit the initial S&P500 futures curve.

In the present case \( \sigma_v(T) \) is a deterministic function of time to maturity, but further development could include either a correlated stochastic process for the volatility, a deterministic local volatility function across moneyness and time to maturity or the use of a stochastic jump similarly to Cont and Kokholm (2013). This will be discussed in further details in section 7.

5.4 VIX futures VaR

In order to calculate the Value-at-Risk (VaR) on the VIX futures or any portfolio consisting of a combination of these contracts, we now have to make some
minor adjustments to our model. When modelling for option valuation, we considered the risk-neutral distribution which implies that the option can be hedged via replication, hence the expected value of the futures contract $VIX_t(T)$ is a martingale and explaining the integration of the forwards variance curve from 0 to $T$ at any time $t$.

We now make adjustments in such ways that we do not suppose that the contract can be hedged hence taking into account the aforementioned roll-down effect and convergence towards the spot. This adjustment will yield in most cases a negative mean return when the VIX curve is in its normal contango shape. This yields the following modification to the model,

\[
d\ln[VIX_t(T - t)] = \ln\left[\frac{v_0(x)}{2} - \int_0^{T-t} \frac{1}{8} \sigma_{VIX^2}^2(T) \sum_{j=1}^{d} \lambda_j (\Phi_j^2)_{xz} dt ds_2 \right. \\
+ \left. \int_0^{T-t} \sum_{j=1}^{d} \sigma_v(T) \sqrt{\lambda_j (\Phi_j^2)_{xz}} dW_{t,s_2} \right] \\
\]

for $t \leq T$.

In order to capture the roll-down effect properly, we also need to model the spot VIX along with the futures. The spot VIX volatility is estimated by fitting the volatility of volatility curve with the same volatility specification used by Cont and Kokholm (2013) which is given by

\[
\sigma_{VIX}(T) = \omega e^{kt} \\
\]

Where we solve numerically for $\omega$ and $k$ to minimise the squared error and then proceed to extrapolate for spot volatility.

6 Results

This section presents the results starting with the model validation, then a Value-at-Risk (VaR) analysis will be done on VIX futures portfolios.

6.1 Model Validation

To validate the model, we must ensure that simulated the forward variance $v_t(T)$ as well as the square-root of the simulated variance swap $VIX_t(T)$ remain martingales. The mean, volatility, and correlations of the simulated data for each maturity on the curve should match what is given in input at any time $t$. The model should also be able to price European VIX options correctly.
6.1.1 Input vs. Output

The first step to validate the model is to take a look at the simulated data against the inputs. The mean, standard deviation and correlations of the simulated data should match the inputs. We simulate the VIX curve for \( t = 1 \) to 90 days with time steps \( dt \) of one day. The test is done with 250,000 simulations using quadratic re-sampling as a variance reduction method. The results in figure 8 and 9 show the simulated VIX futures curve and simulated VIX futures curve volatility against the model inputs at day 90. The absolute relative error on the mean that is found is less than 0.01\% for any maturity and the error on each maturity’s volatility is less than 1\% when comparing to the input volatility.

The statistics of the error on the correlation of the simulated forward variance are presented in table 2. Since a dimensionality reduction method is used to approximate the dynamics between the simulated forward variance maturities, it is expected to see certain discrepancies. However, the average relative error is approximately 4.1\% while in absolute terms this would be approximately 0.035. The standard deviation of the error on the correlation is 4.63\% which indicates a satisfying approximation considering the significant reduction of computing time. Also taking into account that correlations might change through time and the statistics presented in table 2 are not weighted in function of the difference in time to maturity should help the argument that the PCA approximation is valid.

![Mean of Simulated VIX Futures Curve vs. Actual VIX Futures Curve](image)

Figure 8: Simulated VIX Futures Curve vs. Initial VIX Futures Curve
Table 2: Error on Simulated Forward Variance Correlation Between Maturities

<table>
<thead>
<tr>
<th>Metric</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Relative Error</td>
<td>4.10%</td>
</tr>
<tr>
<td>Mean Absolute Error</td>
<td>0.035</td>
</tr>
<tr>
<td>Standard Deviation (Relative)</td>
<td>4.63%</td>
</tr>
</tbody>
</table>

This table shows statistics of the error on the simulated forward variance correlation between maturities. The average relative error is 4.1%. The average absolute error is 0.035 and the relative standard deviation of the error is 4.63%.

Table 3 shows the computing time along with the explanatory power of the PCA method. The explanatory power indicates how much of a good approximation we are making on the curve movements considering the correlations between maturities and how many factors we use from the PCA. The maximum number of factors is seven since the correlation matrix is in this case 7x7 (Spot VIX plus six other maturities). Using seven factors is the equivalent as simulating the seven correlated maturities by using a more classical approach of the Cholesky factorization for correlated random samples hence why it explains 100% of the curve movements. Using less factors means dimensionality reduction as well as reducing the computing time. However, we should expect a certain error to be introduced. The error is dependent on the correlation between forward variance maturities. If the correlation is strong, i.e. close to 1, then fewer factors are
required for a good approximation. The less maturities are correlated, then more factors are then necessary for a good approximation.

As we can observe, for 25,000 simulated paths over 90 time steps, the computing goes from 3.084 seconds using 7 factors to 2.232 seconds using 3 factors. This represents a computing time economy of approximately 38%. Considering the error that is introduced by the method, which is presented in figures 10 to 15 one may choose to sacrifice precision in order to save computing time.

Table 3: Computing Time vs. Number of Factors used and explanatory power of PCA

<table>
<thead>
<tr>
<th>Factors</th>
<th>CPU time</th>
<th>Explained</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>71.80%</td>
</tr>
<tr>
<td>2</td>
<td>2.232s</td>
<td>87.79%</td>
</tr>
<tr>
<td>3</td>
<td>2.548s</td>
<td>93.26%</td>
</tr>
<tr>
<td>4</td>
<td>2.700s</td>
<td>95.98%</td>
</tr>
<tr>
<td>5</td>
<td>2.952s</td>
<td>97.81%</td>
</tr>
<tr>
<td>6</td>
<td>3.084s</td>
<td>99.11%</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td>100%</td>
</tr>
</tbody>
</table>

This table shows the average computing time for 25,000 simulations for 90 time steps and the explanatory power using a different number of factors from the PCA method to approximate the correlation between each maturity.

Figure 10 to 15 present the error distribution on each maturity for the mean and volatility of the simulated samples while using 3, 4 and 7 factors. We have made 1000 simulations of 25,000 paths and took the distribution of the mean and standard deviation error of the simulated samples at step \( t = 90 \) with time step of \( dt = 1/252 \).

The error on the mean for a simulation run is calculated as follow:

\[
\text{MeanError}(\$) = \mathbb{E}[VIX_t(T)] - VIX_0(T)
\]

\[
\text{MeanError}(\%) = 100 \times \left( \frac{\mathbb{E}[VIX_t(T)] - VIX_0(T)}{VIX_0(T)} \right)
\]

Where \( VIX_t(T) \) is the simulated sample for maturity \( T \) at time step \( t \) and \( VIX_0(T) \) is the initial VIX term structure.

The error on the sample volatility is given by

\[
\text{VolError}(\$) = 100 \times \left( \sqrt{\mathbb{E} \left[ \log \left( \frac{VIX_t(T)}{VIX_{t-1}(T)} \right)^2 \right]} - \text{VIXVOL}_0(T) \right)
\]

\[
\text{VolError}(\%) = 100 \times \frac{\text{VolError}(\$)}{\text{VIXVOL}_0(T)}
\]
The results show that there is no a significant loss of precision on the mean when reducing dimensionality. When looking at the average error on the mean in figures 10, 11 and 12, we can see that the distribution of the error are similar as well as the average error being of the same magnitudes. In all three cases, the maximum error is ±0.02% which represents a range of ±0.005$ which is less than the quoted decimal on the contracts.

The error on the mean is of numerical nature since the underlying is distributed log-normally. The variance of the process affects drift term hence committing an error on the volatility of the simulated sample will cause an error in the drift and therefore affect the mean of the simulated sample.

From equation (93) if the volatility of the simulated sample $\hat{\sigma}_v$ differs from the input volatility $\sigma_v$, the adjustment of the log-normal mean will differ in 

$$E_t \left[ -\frac{1}{8} \int_0^T (\sigma_v^2(s_2) \sum_{j=1}^d \lambda_j(\Phi_j^2)_{s_2} dt ds_2) + \frac{1}{8} \hat{\sigma}_V^2 \text{VIX}_2(T) \right].$$

Figure 10: Average error on the initial VIX curve using 3 factors approximation at $t = 90$ days with 1000 simulations of 25,000 paths with distribution of error per maturity.
Figure 11: Average error on the initial VIX curve using 4 factors approximation at $t = 90$ days with 1000 simulations of 25,000 paths with distribution of error per maturity.

Figure 12: Average error on the initial VIX curve using 7 factors approximation at $t = 90$ days with 1000 simulations of 25,000 paths with distribution of error per maturity.
The effects on dimensionality reduction become more apparent when looking at the volatility of the simulated sample. Figures 13, 14 and 15 show the volatility of the simulated samples. If we compare each maturities error distribution of simulated sample volatility, the distributions seem to have roughly the same variance but when looking at the three and four factors case there seems to be a shift in the mean while the distributions of the seven factors errors are centered on 0. This indicates that we might be able to solve the problem with moments matching methods as proposed in Duan and Simonato (1998), however the procedure might counterbalance the computing time economy. Depending on the explanatory power of the PCA method, a choice may be made on the number factors to use in order to make a compromise between computing time and the needed precision.

Figure 13: Average error on the initial VIX volatility using 3 factors approximation at t = 90 days with 1000 simulations of 25,000 paths with distribution of error per maturity.
Figure 14: Average error on the initial VIX volatility using 4 factors approximation at $t = 90$ days with 1000 simulations of 25,000 paths with distribution of error per maturity.

Figure 15: Average error on the initial VIX volatility using 7 factors at $t = 90$ days with 1000 simulations of 25,000 paths with distribution of error per maturity.
Figure 16 shows an example of the different simulated curve shapes that the multi-factorial model allows. As we can see, the use of multiple factors allows the curve to take different shapes. In a 3 factor model, the first factor is predominantly a parallel shift in the curve where there is an increase or decrease of the same magnitude on all maturities. The second factor causes a twist that allows non parallel shifts in the curve. This has an effect on the slope of the curve. An example of a twist would be a small increase in the short term end of the curve and a large increase in the longer maturities or vice-versa. The third factor allows humps which means that certain maturities may move independently from the curve in a manner that will cause a bump in the curve. The more factors we add, the more freely the curve will be able to move.

The figure also shows the mean of the simulated paths for the VIX futures at time $t = 5$ days. When looking at the initial time $t = 0$ VIX curve, we can see a series of squares that represent the mean of the simulated VIX futures at time $t = 5$ days. We can see that on average the VIX futures will roll down the initial curve and converge at the spot.
6.1.2 European VIX Option Valuation

This section presents European VIX Option Valuation. The intuition behind this exercise is based upon interest-rate literature where the no-arbitrage models are in most cases fitted to vanilla options and then used in order to model more complex or exotic payoffs such as path-dependant options. This exercise will enable us to verify that the model follows the process specifications when comparing to another model making the same assumptions, thus insuring that we are able to price simpler products correctly before using the model in applications that require more flexibility.

In section 5.3, we use a variance swap on the VIX to find the volatility specifications of the forward variance. This measure approximates the average VIX implied volatility across moneyness for a certain maturity and is used in this case to simplify the problem. Instead of using the VIX variance swap as an input for the volatility of the forward variance, we could follow the same reasoning that is shown in section 5.3 and directly back out the forward variance volatility surface from the VIX implied volatility surface.\textsuperscript{13} The model’s options prices should then match exactly market VIX options prices, thus in one case or the other yielding the same options prices as the Whaley model across time to maturity and moneyness since we are making the same assumptions of log-normality of the underlying.\textsuperscript{14} The model implementations allowing to capture the skew of the implied distribution will be further discussed in section 7.

Since we are making the same assumptions on the VIX futures distribution as the Whaley model, we should be able to price European VIX options and obtain the same results for the same VIX volatility inputs. Table 5 shows the results. A set of options of 9, 29, and 53 days to maturity was priced according to different moneyness. Being able to price vanilla options comparatively to the Whaley model should confirm that our model follows the process specifications that we have initially set.

Table 4 shows the volatility inputs for both model, starting with the square-root of the VIX variance swap as volatility specification that is used as input for the Whaley model. Having defined the forward variance correlations, we are then able to find the volatility specifications of the forward variance from equation 92 of section 5.3.

\begin{footnotesize}
\begin{itemize}
  \item \textsuperscript{13}Obtained with the Whaley model.
  \item \textsuperscript{14}In our case the forward variance is log-normal, which makes the integration over time to maturity that yields the VIX futures price log-normal as well.
\end{itemize}
\end{footnotesize}
Table 4: Model Inputs: VIX Futures Volatility, Forward Variance Volatility, and Correlations Between Forward Variance Maturities

<table>
<thead>
<tr>
<th>Maturity</th>
<th>1&lt;sup&gt;st&lt;/sup&gt;</th>
<th>2&lt;sup&gt;nd&lt;/sup&gt;</th>
<th>3&lt;sup&gt;rd&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Days to maturity</td>
<td>9</td>
<td>29</td>
<td>53</td>
</tr>
<tr>
<td>Volatility</td>
<td>1.1049</td>
<td>0.9175</td>
<td>0.7951</td>
</tr>
<tr>
<td>Forward Variance Volatility</td>
<td>2.2098</td>
<td>1.4065</td>
<td>1.2745</td>
</tr>
<tr>
<td>Correlations Between Maturities</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1&lt;sup&gt;st&lt;/sup&gt;</td>
<td>1</td>
<td>0.8493</td>
<td>0.6982</td>
</tr>
<tr>
<td>2&lt;sup&gt;nd&lt;/sup&gt;</td>
<td>0.8493</td>
<td>1</td>
<td>0.8221</td>
</tr>
<tr>
<td>3&lt;sup&gt;rd&lt;/sup&gt;</td>
<td>0.6982</td>
<td>0.8221</td>
<td>1</td>
</tr>
</tbody>
</table>

The table shows the volatility inputs, starting with the square-root of the VIX variance swap as volatility specification that is used as input for the Whaley model. Having defined the forward variance correlations, we are then able to find the volatility specifications of the forward variance from equation 92 of section 5.3.

The results are presented in table 5 and compare 100,000 Monte-Carlo simulations for the forward-variance model against the Whaley closed-form. The presented moneyness in the table is calculated from \( \ln(K/S_0) \). \( S_0 \) being the spot VIX (22.97) and \( K \) the strike price of the option.

As expected, the results indicates that the two models give very similar results for the same VIX futures volatility inputs. The error between the two models across days to maturity and moneyness is in most cases less than 1%. In any case the absolute error is significantly smaller than the bid-ask spread of the traded VIX options. This error might simply be caused by numerical error of the Monte-Carlo simulation. Adding more simulations or using variance reduction methods should yield more precise results. This shows however that by taking the volatility parameters directly from the VIX options implied volatility, we should then be able obtain the correct market prices across moneyness and time to maturity for the VIX options with a relatively small error.
<table>
<thead>
<tr>
<th>Days to Maturity</th>
<th>Model</th>
<th>Volatility</th>
<th>Moneyness</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>-0.19</td>
<td>-0.14</td>
</tr>
<tr>
<td>9</td>
<td>Whaley</td>
<td>1.1049</td>
<td>5.99</td>
</tr>
<tr>
<td>29</td>
<td></td>
<td>0.9167</td>
<td>7.36</td>
</tr>
<tr>
<td>53</td>
<td></td>
<td>0.7951</td>
<td>8.55</td>
</tr>
<tr>
<td>9</td>
<td>Forward</td>
<td>2.2098</td>
<td>5.97</td>
</tr>
<tr>
<td>29</td>
<td>Variance</td>
<td>1.4065</td>
<td>7.36</td>
</tr>
<tr>
<td>53</td>
<td>Simulation</td>
<td>1.2745</td>
<td>8.57</td>
</tr>
<tr>
<td>Absolute Error</td>
<td>9</td>
<td>-0.0146</td>
<td>-0.0158</td>
</tr>
<tr>
<td></td>
<td>29</td>
<td>-0.0023</td>
<td>-0.0031</td>
</tr>
<tr>
<td></td>
<td>53</td>
<td>0.0143</td>
<td>0.0142</td>
</tr>
<tr>
<td>Relative Error</td>
<td>9</td>
<td>-0.243%</td>
<td>-0.308%</td>
</tr>
<tr>
<td></td>
<td>29</td>
<td>-0.031%</td>
<td>-0.048%</td>
</tr>
<tr>
<td></td>
<td>53</td>
<td>0.168%</td>
<td>0.185%</td>
</tr>
</tbody>
</table>

The table shows the comparison on European option pricing between the Whaley model (closed form) and the simulated forward variance model (100,000 simulations) for different maturities and moneyness. Moneyness is calculated from $\ln\left(\frac{K}{S_0}\right)$, where $S_0$ is the spot VIX (22.97) and $K$ the strike price of the option. Using the same inputs for both models, the results show an error that is relatively small (>1% on average) and in any case significantly smaller than the usual bid-ask spread on these options.
6.2 Distributions and Value-at-Risk

This section presents results for the distributions and VaR given by the model from equation 94 and compare two variations. One of which uses three factors against one that uses only one factor (perfectly correlated maturities). We will first take a look at the distributions and VaR of a long and a short position of one contract\textsuperscript{15} in a VIX futures. We then compare the distributions and VaR on a VIX futures portfolio. The evaluation date is January 3\textsuperscript{rd} 2012. The initial VIX term structure details are given in table 6. Implied volatility is calculated on the day of the evaluation with the square-root of the variance swap on the VIX options as in the VIX methodology in equation 2.

\begin{table}[h]
\centering
\begin{tabular}{lcccccc}
\hline
Maturity & Spot & 1\textsuperscript{st} & 2\textsuperscript{nd} & 3\textsuperscript{rd} & 4\textsuperscript{th} & 5\textsuperscript{th} & 6\textsuperscript{th} \\
Days to maturity & 9 & 29 & 53 & 72 & 92 & 116 \\
\hline
Price & 22.97 & 24.77 & 25.77 & 26.80 & 27.48 & 27.82 & 28.20 \\
Implied Volatility & 1.2910 & 1.1049 & 0.9175 & 0.7951 & 0.7120 & 0.6552 & 0.6040 \\
\hline
\end{tabular}
\caption{VIX term structure on January 3rd 2012}
\end{table}

Let’s begin with a position in a single futures contract. We will first take a look at a long position in the 2\textsuperscript{nd} maturity contract which has 29 business days to expiration. The analysis is made at $t = 5$ days. Figure 17 and 18 present the distributions and VaR obtained with the one factor model and the three factor model respectively. Table 7 presents each distribution’s parameters for the two model variations. As we are able to see in table 7, the use of multiple factors has little impact on the distribution parameters where all four presented moments show slight differences. The VaR measures at 95\% and 99\% are substantially the same for both model variations with no significant differences. The next case in figures 19 and 20 present the distribution of a short position in the same contract. The distributions parameters remain the same as for the long position but the VaR changes. Once again the VaR between the two model variations almost shows no significant differences.

\begin{table}[h]
\centering
\begin{tabular}{lcccc}
\hline
Model & Mean & Standard Dev. & Skewness & Kurtosis \\
\hline
1 factor & -251.45 & 3488.52 & 0.5108 & 3.4544 \\
3 factors & -250.74 & 3443.43 & 0.4873 & 3.4373 \\
\hline
\end{tabular}
\caption{Distribution parameters on 2\textsuperscript{nd} maturity contract P&L}
\end{table}

\textsuperscript{15}VIX futures contracts have a multiplier of 1000 meaning that a change of 0.01 in the quoted price is equal to 10$. 

63
Figure 17: 1 factor model P&L Distribution of on a long position consisting of one contract of the 2nd maturity (29 business days to expiration) VIX futures at $t = 5$ days along with 95% and 99% VaR estimated with the model from equation 94 with 100,000 simulations

Figure 18: 3 factors model P&L Distribution of on a long position consisting of one contract of the 2nd maturity (29 business days to expiration) VIX futures at $t = 5$ days along with 95% and 99% VaR estimated with the model from equation 94 with 100,000 simulations
Distribution of 5 Days P&L for a Short Position in the 2\textsuperscript{nd} Maturity Contract

VaR 95\% = $-5938.35$

VaR 99\% = $-9085.83$

Mean P&L = $251.45$

Figure 19: 1 factor model P&L Distribution of on a short position consisting of one contract of the 2\textsuperscript{nd} maturity (29 business days to expiration) VIX futures at $t = 5$ days along with 95\% and 99\% VaR estimated with the model from equation 94 with 100,000 simulations

Distribution of 5 Days P&L for a Short Position in the 2\textsuperscript{nd} Maturity Contract

VaR 95\% = $-5826.7$

VaR 99\% = $-8953.16$

Mean P&L = $250.74$

Figure 20: 3 factors model P&L Distribution of on a short position consisting of one contract of the 2\textsuperscript{nd} maturity (29 business days to expiration) VIX futures at $t = 5$ days along with 95\% and 99\% VaR estimated with the model from equation 94 with 100,000 simulations
Although the distributions for a position in one contract are undeniably similar from one variation of the model to another, let’s take a look at the distribution of the roll-down effect for both variations of the model. Figures 21 and 22 show the distribution obtained with the one factor model and three factors model respectively. Despite the fact that both means are very similar, the shape of the distributions and higher moments differ significantly. From the descriptive statistics in table 8 and the shapes of the distributions, we can make the statement that the one factor model is in fact limiting when it comes to the number of possible scenarios.

Table 8: Distribution parameters on Roll Down Effect in 2\textsuperscript{nd} maturity contract

<table>
<thead>
<tr>
<th>Model</th>
<th>Mean</th>
<th>Standard Dev.</th>
<th>Skewness</th>
<th>Kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Factor</td>
<td>-251.45</td>
<td>136.04</td>
<td>1.0916</td>
<td>4.9722</td>
</tr>
<tr>
<td>3 Factors</td>
<td>-250.74</td>
<td>166.16</td>
<td>0.6361</td>
<td>3.9215</td>
</tr>
</tbody>
</table>

Figure 21: 1 factor roll down effect P&L distribution on the 2\textsuperscript{nd} maturity (29 business days to expiration) VIX futures at $t = 5$ days with 100,000 simulations
Following on the fact that the number of factors has an impact on the roll down distribution, this explains the slight differences in the single contract distributions and should consequently have a more pronounced effect on a portfolio holding multiple contracts of different maturities.

Let's now consider a portfolio that is short one futures contract of the first maturity and long one contract of the fourth maturity with the market conditions in Table 6.

Following earlier discussions and empirical evidence from Whaley (2013), this strategy should have a positive mean return caused by the roll-down effect since the VIX curve is currently in contango and the slope is steeper between the 1st month and the spot than between the 3rd and 4th months. Distributions for both model variations methods are presented in Figure 23 and 24 along with the parameters of each distribution in Table 9.

Once more, even with similar means Table 9 shows that the higher moments differ between both distributions.

<table>
<thead>
<tr>
<th>Model</th>
<th>Mean</th>
<th>Standard Dev.</th>
<th>Skewness</th>
<th>Kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Factor</td>
<td>784.35</td>
<td>1239.13</td>
<td>-0.9840</td>
<td>4.6133</td>
</tr>
<tr>
<td>3 Factors</td>
<td>778.98</td>
<td>1431.21</td>
<td>-0.6452</td>
<td>3.8756</td>
</tr>
</tbody>
</table>
Figure 23: 1 Factor P&L distribution calendar spread strategy between 1st and 4th maturity VIX futures at $t = 5$ days along with 95% and 99% VaR estimated with the model from equation 94 with 100,000 simulations.

Figure 24: 3 Factor P&L distribution for the calendar spread strategy between 1st and 4th maturity VIX futures at $t = 5$ days along with 95% and 99% VaR estimated with the model from equation 94 with 100,000 simulations.
When comparing both distributions, we can see that the three factors model enables more possible outcomes on both the up and down side. The VaR for the one factor model is undervalued at 95% and 99% when comparing to the three factors model VaR.

Figure 25: 3 Factor and 1 factor P&L cumulative distribution for the calendar spread strategy between 1st and 4th maturity VIX futures at $t = 5$ days with 100,000 simulations.

Figure 25 shows the cumulative distribution obtained from both model variations. Again we are able to see that more probabilities are assigned to extreme events with the use of three factors as well as more probabilities assigned to loss events altogether. The indicated loss probability ($P\&L < 0$) with the one factor model is 0.2261 while being 0.2665 when using three factors.

## 7 Discussion and further improvements

The shape of the term structure has a direct impact on the VIX futures returns, it is important to take it into account. By using the no-arbitrage framework, we are able to build the VIX curve dynamics that relies on actual given market data instead of parameter calibration but also fit the initial term-structure perfectly in any given situation. The use of the three Brownian motions driving the curve also play an important role in allowing the simulated term structure to take different shapes. Our results show that the number factors used to drive the curve dynamics has a significant difference for the analysis of a VIX futures portfolio of different maturities.
Taking in consideration previous research, the presented model seems a more realistic approach for risk management when dealing with a portfolio of VIX futures. Since we are dealing with a forward looking measure, the usefulness of historical data is arguable when estimating distribution parameters. We can however rely on market information to build the distributions of VIX futures or price other VIX derivatives. As previously mentioned in section 5.4, the relation between the VIX realized and implied volatility should be studied in order to determine if the VIX implied volatility is in fact an efficient predictor of its realized volatility as it is the case with stock indexes.

Aside from obtaining a more realistic VIX futures VaR model, the model could also be used to price VIX exotic options or VIX ETPs with their respective options. Like the VIX, the VXX (or any other VIX ETP) has a constant time to maturity. This fixed time to maturity can however change from one rolling period to the other depending on the number of business days between the first and second month contract. Usual time to maturity ranges between 18 and 25 business days. When evaluating the VXX from the implied volatility curve, its maturity will be seen as 21 business days\(^{16}\) (VIX time to maturity) plus the time to maturity between the first and second VIX contracts. Let’s define \(\tau(t)\) as the number of business days between the first and second VIX contracts. The expression of the VXX in terms of VIX is then given by.

\[
VXX_t = VXX_{t-1} \times \left[ 1 + TBR_t + \left( \frac{VIX_t(T + \tau(t))}{VIX_{t-1}(T + \tau(t))} \right) \right] (98)
\]

Where \(T\) is 21 business days and \(VIX_t(T + \tau(t))\) is found by using a weighted linear combination of the two first active contracts as in equation (17).

To find the call price, we apply the payoff \(\max(VXX_t - K, 0)\) for a call option and \(\max(K - VXX_t, 0)\) for a put option. The call price is then given by

\[
C_{VXX} = e^{-rt} \mathbb{E}^Q[\max(VXX_t - K, 0)] (99)
\]

and the put price by

\[
P_{VXX} = e^{-rt} \mathbb{E}^Q[\max(K - VXX_t, 0)] (100)
\]

It should be important to note that the roll-down effect previously discussed doesn’t have any impact on the option prices since the valuation is within a risk-neutral framework. We are merely evaluating a VIX futures option with a defined maturity \((T + \tau(t))\), thus the \(VXX_t\) price from the model should be a martingale i.e. \(\mathbb{E}^Q[VXX_t] = VXX_0\).

More sophisticated modifications could also be looked upon in order to obtain an implied distributions of the VIX futures closer to the market expectations. For example, we have overlooked the skewness aspect of the distribution by using a deterministic volatility of volatility that is only a function of time to

\(^{16}\)30 calendar days \(\approx\) 21 business days.
maturity. By using a volatility surface in the model instead of a variance swap term-structure for the volatility of the forward volatility, we should be able to capture the positive correlation between the VIX and its volatility and therefore the skewness of the distribution. Staying consistent with the market model approach, this could be achieved by using local volatility as in Dupire (1997). The approach can however become tricky since we would have to worry about strike arbitrage in addition to calendar arbitrage on VIX options as well as calendar arbitrage on the VIX term-structure and would have to undergo certain modifications since we would have to back out the local volatility of the forward variance from the implied volatility surface of the VIX options. Another solution to capture the positive skewness of the VIX distribution would be to include a stochastic jump in the model as in Cont and Kokholm (2013), however model calibrations would be needed to have the proper parameters for the Poisson process used for the jump in order to fit the implied volatility of the VIX futures.

8 Conclusion

From previous literature on no-arbitrage volatility models, we have used a forward variance model to obtain a VIX futures term-structure process that is free of arbitrage. Knowing from Whaley (2013) that the slope of the VIX futures term-structure plays an important role in predicting their returns, we have argued that the no-arbitrage framework should be better suited to capture this measure in risk analysis. By taking the term structure as an input, the model fits the initial term structure perfectly and the use of principal component analysis for the curve dynamics allows it to take different shapes.

Our results show that when comparing our multi-factorial model to a one factor model, there is significant difference in the distribution of the roll-down effect of the VIX futures contract. The differences in the roll-down distribution when using a multi-factorial model has a slight impact on the skewness and kurtosis of the distribution when holding a position in a single VIX futures contract but becomes far more apparent when holding a VIX futures portfolio of different maturities. This in turn impacts significantly the Value-at-Risk (VaR) of a portfolio holding multiple maturities.

The results also show that the use of PCA to model the curve’s dynamics reduces the number of factor needed to drive the curve while still capturing most of the possible curve scenarios (93% vs. 71% when using only one factor). When comparing to more classical methods such as correlated random samples generation using Cholesky factorisation that would allow to capture 100% of the possible curve’s movements but where we would need one factor per simulated maturity, the PCA reduces significantly the computing time (approximately 38%) when simulating the entire VIX curve that usually includes the spot VIX along with six additional forward maturities.
The model can be used jointly with a process for the S&P500 that gives a complete market free of arbitrage and can also be used to price VIX exotic options as well as VIX ETPs options.

Modifications are still to come regarding the volatility of volatility which in this case was a deterministic function of time to maturity calculated by the square-root of a variance swap from the VIX European options. Further developments should include a skewness component that should be added by taking into account the volatility across moneyness on the VIX options. This in turn would allow to capture the correlation between the volatility and the volatility of volatility.
References


