The fear gauge

VIX volatility index and the time-varying relationship between implied volatility and stock returns

Anssi Heinonen

University of Helsinki
Faculty of Social Sciences
Economics

Master's Thesis

May 2013

Tiedekunta/Osasto – Fakultet/Sektion – Faculty Faculty of Social Sciences	Laitos – Institution – Department Department of Political and Economic Studies	
Tekijä – Författare – Author Anssi Heikki Heinonen		
Työn nimi – Arbetets titel – Title The fear gauge – VIX volatility index and the time-varying relationship between implied volatility and stock returns		
Oppiaine – Läroämne – Subject Economics		
Työn laji – Arbetets art – Level Master's thesis	Aika – Datum – Month and year May 2013	Sivumäärä – Sidoantal – Number of pages 72

Tiivistelmä – Referat – Abstract

Implied volatility is the level of dispersion of asset price changes that is embedded in the market prices of option contracts written on that asset. As such, it represents market participants' consensus on the expected volatility, or uncertainty regarding future returns, over the remaining life of the options. A proxy for aggregate stock market implied volatility is given by Chicago Board Options Exchange's volatility index, the VIX, which is a practical implementation of the concept of model-free implied volatility that allows for extracting implied volatility directly from observed option prices without the use of a parametric option pricing model. Two well-known features of implied volatility and the VIX in particular are its mean reversion – it tends towards its mean level over time – and a significant negative contemporaneous correlation with returns on the underlying stock index. In this thesis, the time-varying dynamics of the relationship between stock market returns and implied volatility are examined empirically.

Theoretical background and the construction methodology of the VIX are thoroughly discussed in order to give the reader an understanding of model-free implied volatility and the VIX. Structural stability of the time series data of the VIX daily levels ranging from January 2004 to December 2011 is tested using the method developed by Bai and Perron (1998;2003a;b;2004) for testing for structural breaks at a priori unknown times. The results suggest that statistically significant structural breaks are present in the data sample, that is, the mean level to which the VIX reverts is found to shift over time. This allows for determining distinct volatility regimes in the sample. The relationship between daily changes in the VIX and the corresponding daily stock index returns during these distinct regimes is studied by examining their cross-correlations and testing for any Granger causality that might exist. A more elaborate study is conducted on the strongly negative contemporaneous relationship by examining the effect of return shocks on volatility as well as by looking for asymmetries in volatility using linear regression.

The results strongly suggest the presence of a single statistically significant structural break that coincides approximately with the outbreak of the global financial crisis in late 2007. The sample data is therefore divided into a pre-breakpoint low volatility regime and a post-breakpoint high volatility regime. No statistically significant Granger causality is found in the data, which suggests that VIX changes have no predictive power over stock index returns or vice versa. Closer scrutiny of the contemporaneous volatility-return relationship reveals asymmetries in volatility – increases in the VIX associated with negative stock market returns are found to be higher than decreases associated with positive returns of similar size. The overall inverse relationship appears to be stronger in the high volatility regime. However, the degree of asymmetry in that relationship is in turn stronger in the low volatility regime, i.e. the difference between increases and decreases in the VIX in response to negative and positive returns, respectively, turns out to be higher than in the high volatility regime.

The study of structural stability of the VIX mean level provides and update to the study of Guo and Wohar (2006), whose sample period predates that of this thesis. The findings on the asymmetry of the relationship between implied volatility and stock market returns lend further support to the notion that the VIX is rather a measure of investor fear in falling markets than of investor positive sentiment in rising markets, which has earned it the moniker "investor fear gauge". In a low volatility regime, investors are more sensitive to any decreases in the stock markets, whereas in a high volatility regime the effects of return shocks are more pronounced regardless of their sign. The VIX constitutes a powerful indicator of investor sentiment, i.e. the expected level of volatility perceived by market participants at any given time.

Avainsanat – Nyckelord – Keywords
Volatility index
Model-free implied volatility
Structural breaks
Stock returns

Contents

1 Introduction **5**

2 Pricing of financial derivatives **8**

2.1 Definitions 8

2.2 Asset price processes 9

2.3 Risk-neutral valuation of contingent claims 12

3 Gauging volatility **16**

3.1 Model-free implied volatility 17

3.2 Variance swap as model-free implied volatility 20

3.3 The VIX volatility index 23

4 Volatility regimes: Identifying structural breaks in the VIX time series **29**

4.1 Normal ranges 31

4.2 Econometric methodology 32

4.3 Results 37

5 Relationships between implied volatility and stock market returns **40**

5.1 Summary statistics 41

5.2 Intertemporal cross-correlations 45

5.3 Granger causality tests 47

5.4 Asymmetries in volatility 50

6 Conclusions **58**

References **61**

A **68**

A.1 The equivalence of model-free implied variance and the fair value of future variance . . . 68

A.2 Portfolio replication strategy . . . 70

List of Figures

4.1 VIX daily time series from January 2004 to December 2011. 30
4.2 Volatility regimes and their mean levels in the VIX daily time series from January 2004 to December 2011. 39

5.1 Daily closing values of the VIX volatility index (left scale) and the S&P 500 stock index (right scale) from January 2004 to December 2011. 45
5.2 Sample cross correlations between lagged SPX returns and changes in the VIX in Regimes 1 and 2. 47
5.3 News impact curves for the VIX in the full sample and in Regimes 1 and 2. 53

List of Tables

4.1 Yearly percentile ranges for VIX daily closing levels from January 2, 2004 to December 31, 2011. 31
4.2 Bai and Perron structural break test results for the VIX series, sample period 02/01/2004–31/12/2011. 38

5.1 Descriptive statistics of the VIX and the S&P 500 levels and returns for the full sample period and Regimes 1 and 2. 43
5.2 Results of Granger causality tests for SPX returns and VIX changes. 50
5.3 Impact of return shocks on implied volatility. 54
5.4 Regression results for Models (5.10) and (5.11). 57

Chapter 1

Introduction

Volatility is undoubtedly one of the key variables in finance. It is defined conventionally as the variability or dispersion of asset price changes over time – as a statistical measure, it corresponds to the standard deviation (square root of variance) of asset returns and is therefore usually denoted by $\sigma(\cdot)$. In the context of pricing financial derivatives, volatility is perhaps the most important variable; to price an option contract, for example, one must estimate the volatility of the asset on which the option is written, over a future period from now until the expiry of the option contract. Conversely, the value of volatility derived from market prices of traded derivatives yields a measure that reflects future expectations of all market participants. This *implied volatility* constitutes a forward-looking estimate for the dispersion of returns on the underlying asset. On an aggregate market level, the implied volatility of e.g. an equity index that is viewed as a proxy of general market conditions could then be thought of as a measure of perceived uncertainty in that market.

Volatility is readily estimated from historical data using sample standard deviation, but defining implied volatility can prove to be trickier. Traditionally, it is estimated by inverting some option pricing formula, but more recent research has led to the development of an informationally superior method that allows extracting volatility from option prices without the need to specify an option pricing model. A practical implementation of this model-free measure of implied volatility is found in the Chicago Board Options Exchange's VIX volatility index that aims to capture the market's expectation of future volatility over the next 30 calendar days. The VIX is often informally referred to as the "investor fear gauge", a term coined by Whaley (2000) to underline the observation that the volatility index increases in times of declining markets, thus reflecting market stress and investors' fear of market crashes. More formally, there is a strong inverse relationship between changes in the volatility index level and corresponding returns on the underlying stock index.

A well-known feature of market volatility is that it is a mean-reverting process, which essentially means it will tend to its average level, increasing (decreasing) from low (high) levels over time. However, the mean level to which market volatility reverts is likely to shift over time, reflecting a transition from a regime of high or low volatility to another. Indeed, graphical inspection of VIX daily time series data reveals infrequent, but significant shifts in its mean level. In recent years, market conditions have alternated from relatively calm rising markets to extremely volatile, at times even chaotic periods witnessed after the global financial crisis first began in late 2007. Obviously, the market turbulence is quantified in not only the observed level of market volatility but also stock market returns. As such, the degree and dynamics of their interrelationship is expected to vary accordingly. For example, during favorable market conditions investors might be more wary of any negative asset returns, amplifying their reactions and thus causing heightened levels of implied volatility compared to a murkier market environment. We are therefore interested in ascertaining statistically whether there have been any significant structural breaks in the VIX mean level and in identifying the dates of their occurrence. These breakpoints can then be used in defining the boundaries between different volatility regimes, or distinct subperiods during which the supposedly changing dynamics of the relationship between market volatility and asset returns can be examined.

This thesis contributes to the existing market volatility literature in at least two ways. First, we provide a thorough, yet concise synthesis of the theoretical framework underlying the VIX index. This should be sufficient for the unacquainted reader to get a grasp on what the VIX is and how it is constructed. The new calculation methodology no longer relies on the widely-used Black-Scholes (1973) model. Instead, it is based on the concept of fair value of future variance developed by Demeterfi, Derman, Kamal, and Zou (1999) and calculated directly from option prices independent of any model. The theory is closely linked to the pricing of variance swaps, whose fair strike price is shown to be identical to the separately developed concept of model-free implied variance, or squared volatility (Britten-Jones and Neuberger 2000).

Second, we examine the empirical properties of the VIX, namely structural stability and the relationship with stock market returns, over a sample period that covers the times of extraordinary market turbulence experienced in recent years. We extend the study of Guo and Wohar (2006) on structural changes in market volatility by examining more recent data and find a regime shift that coincides with the outbreak of the global financial crisis in late 2007. This leads to the division of the sample data into two volatility regimes with a significant difference between their mean levels of volatility. Moreover, we aim to provide an update to existing studies covering the relationship between implied volatility, as measured by the VIX, and the corresponding asset returns. Historically, the VIX exhibits a strong negative correlation with returns on its underlying stock index, the S&P 500. This inverse relationship has merited the VIX its nickname, the "fear gauge": The index value tends to increase during times of financial stress, which are often accompanied by market decline, i.e. negative asset returns. Conversely, in a rising market the VIX usually goes down. Our findings give strong evidence of an inverse relationship and of asymmetries between negative and positive changes in volatility vis-à-vis corresponding asset returns. Interestingly, the strength of the inverse relationship as well as the degree of asymmetry therein are found to vary between the two volatility regimes.

The remainder of this paper is organized as follows. Chapter 2 presents the terminology, concepts and key results in mathematical finance and option pricing insofar as they are necessary for understading model-free implied volatility introduced in the next chapter. The first section of Chapter 3 begins with a brief discussion on financial volatility before moving on to examine the theory of model-free implied volatility in the next section. The practical implementation and construction of the VIX volatility index is then presented in the third section. Chapters 4 and 5 constitute the empirical part of this study. In Chapter 4, we look for structural breaks in the VIX time series and, based on the results, divide the sample period into two volatility regimes. Chapter 5 studies the dynamics between contemporaneous changes in the VIX and the underlying stock index returns in the context of the newly defined volatility regimes. Chapter 6 concludes. Some auxiliary mathematical proofs are contained in the Appendix.

Chapter 2

Pricing of financial derivatives

In this chapter we introduce the concepts, terminology and mathematical foundations of pricing contingent claims, insofar as they are relevant for the scope of this thesis. Our presentation draws partly from Hull (2009), Tsay (2005) and Øksendal (1998) and is intended to give the reader the tools necessary for understanding the concept of financial volatility in the context introduced later on. For a more rigorous treatment of stochastic calculus and the underlying theories, the interested reader is referred to e.g. Neftci (2000) and Øksendal (1998).

2.1 Definitions

In financial markets, a *derivative* is an instrument whose value is contingent on another variable known as the *underlying asset*. Such assets can be e.g. stocks, currency exchange rates or commodities.

Perhaps the most simple derivative is a *forward contract*. It is an agreement between two parties to buy or sell an underlying asset at a future time specified today and at a price agreed upon today. The party who buys the underlying asset is said to assume a *long position*, whereas the seller assumes a *short position*. The price agreed upon is called the *delivery* or (especially in conjunction with options) *strike* price and denoted K, and the future trade date is called the *expiry* or *maturity date* T. At expiry, the forward contract gives a payoff of $S_T - K$ for a long position and $K - S_T$ for a short position, where S_T is the price of the underlying asset at time T.

An *option* is a contract which gives the holder the right, but not the obligation, to buy or sell an underlying asset at a later time for a strike price that is fixed when the option is written. If the holder has the right to buy the underlying asset, the option is a *call*, whereas an option that gives its holder a right to sell the underlying asset is a *put*. The use of this right is called *exercise* of the option. If an option can only be exercised at the expiry date T it is called *European*, while exercising an *American* option is allowed at any time before the expiry date. Rational investors will exercise an option only when it is profitable for them; in the case of a call option, this happens when the market price of the underlying asset at any given time t is greater than the strike price, i.e. $S_t > K$. Correspondingly, a put option is only exercised when $S_t > K$. The difference between the two prices gives the *payoff* of the option. Denoting the payoff function of a call and a put option at time t with C_t and P_t, respectively, gives

$$C_t = \max(S_t - K, 0) \equiv (S_t - K)^+ \tag{2.1}$$

$$P_t = \max(K - S_t, 0) \equiv (K - S_t)^+. \tag{2.2}$$

2.2 Asset price processes

A variable whose value changes over time in an uncertain way is said to follow a *stochastic process*. The process can be discrete or continuous in time, depending on whether the changes in value take place at fixed points in time or at any given time, as well as in "space", depending on whether the variables of the process can take only certain values, or all values within a given range. The birth of modern mathematical finance can be attributed to Bachelier (1900), who proposed modeling the price process $\{S_t\}_{t\geq 0}$ of a financial asset as a stochastic process in order to develop a theory of option pricing. In spite of the fact that in financial markets the prices follow processes that are discrete in time and have discrete variables, since the prices are quoted in fixed ticks and trading can only take place when the marketplace is open for business, continuous-time processes are still useful for understanding the pricing of options.

A (first-order) *Markov* process is a certain stochastic process where only the present value of a stochastic variable is relevant when considering the next value. This implies that any past information W_j with $j < t$ is irrelevant when forecasting the future values. A *Wiener process*[1] is a particular Markov process with mean 0 and variance 1. Denote such

[1] Also referred to as *Brownian motion*, particularly in physics. A discrete-time analogue would be a white noise process.

a process with W_t. Its changes $dW = W_{t+dt} - W_t$, where d is an infinitesimal change, are then given by

$$dW_t = \epsilon\sqrt{dt}. \tag{2.3}$$

Here, ϵ is a random drawing from a standardized normal probability distribution; therefore, $dW \sim N(0, dt)$. The Wiener process W_t with $W_0 = 0$ is further characterized by

$$W_t - W_s \sim N(0, t-s)\ \forall\ 0 \leq s \leq t \text{ and}$$
$$W_{t_1} - W_{s_1} \perp\!\!\!\perp W_{t_2} - W_{s_2}\ \forall\ s_1 \leq t_1 \leq s_2 \leq t_2,$$

where the latter condition implied that two non-overlapping time intervals are independent. From the former condition it is evident that $E[W_t] = 0$ and $\mathrm{Var}[W_t] = t$, i.e. the variance increases linearly with the time interval. A *generalized Wiener process* is a generalization where the rates of change of its mean and variance are allowed to differ, so that

$$dx_t = \mu dt + \sigma dW_t, \tag{2.4}$$

where μ is the *drift rate* (change in mean) and σ is the rate of change in the standard deviation (square root of variance), henceforth referred to as *volatility*[2]. Thus, dx_t is normally distributed with mean μt and variance $\sigma^2 t$.

A further extension known as the *Itō process*[3], is one where the drift μ and volatility σ are functions of the stochastic process itself, that is,

$$dx_t = \mu(x_t, t)dt + \sigma(x_t, t)dW_t, \tag{2.5}$$

meaning that both the drift and volatility parameters are allowed to vary through time.[4] The prices of financial assets are generally assumed to follow this process, which makes it essential for financial modeling. For example, let S_t be an asset whose price process $\{S_t\}_{t\geq 0}$ is characterized by an Itō process with $\mu(x_t, t) = \mu S_t$ and $\sigma(x_t, t) = \sigma S_t$:

$$dS_t = \mu S_t dt + \sigma S_t dW_t \Leftrightarrow \frac{dS_t}{S_t} = \mu dt + \sigma dW_t. \tag{2.6}$$

This particular model specification is known as *geometric Brownian motion*. The latter

[2]Financial volatility is discussed with more detail in the following chapter.
[3]Named after Kiyoshi Itō for his work on stochastic calculus (see e.g. Itō 1951).
[4]Note that the Wiener process derived above is simply an Itō process with $\mu(x_t, t) = 0$ and $\sigma(x_t, t) = 1$.

expression states that the underlying asset's price *return* process follows a generalized Wiener process, which allows for some convenient properties on the evolution of the return process.

Another result that is employed widely in mathematical finance and also carries Itō's name is *Itō's lemma*. More generally, it is often perceived as one of the cornerstones of stochastic calculus – it is the stochastic counterpart of the chain rule.[5] In its simplest (one-dimensional) form it states that for an Itō process given by equation (2.5) and for any contingent twice differentiable function $f(x_t, t)$ one has

$$df = \left(\frac{\partial f}{\partial t} + \frac{\partial f}{\partial x}\mu(x_t, t) + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}\sigma^2(x_t, t) \right) dt + \frac{\partial f}{\partial x}\sigma(x_t, t) dW_t. \tag{2.7}$$

A useful application is to consider Itō's lemma with the logarithm of the asset price S_t that follows a geometric Brownian motion in order to arrive at an analytical solution for S_t. Let $x_t = S_t$ and $f(x_t, t) = \ln S_t$, so that equation (2.7) becomes

$$\begin{aligned} d\ln S_t &= \left(\frac{1}{S_t}\mu S_t - \frac{1}{2}\frac{1}{S_t^2}\sigma^2 S_t^2 \right) dt + \frac{1}{S_t}\sigma S_t dW_t \\ &= \left(\mu - \frac{\sigma^2}{2} \right) dt + \sigma dW_t, \end{aligned} \tag{2.8}$$

By integrating equation (2.8), we have

$$\begin{aligned} \ln S_t - \ln S_0 &= \left(\mu - \frac{\sigma^2}{2} \right) t + \sigma W_t \\ S_t &= S_0 \exp\left[\left(\mu - \frac{\sigma^2}{2} \right) t + \sigma W_t \right]. \end{aligned} \tag{2.9}$$

Moreover, using the expected value of a Gaussian (normally distributed) random vari-

[5]For a composite function $f \circ g$, the chain rule is given by

$$\frac{df}{dx} = \frac{df}{dg}\frac{dg}{dx}.$$

able, $E[e^X] = \exp(\mu + \frac{\sigma^2}{2})$, the expected asset price can be shown to equal

$$E[S_t] = E\left[S_0 e^{\left(\mu - \frac{\sigma^2}{2}\right)t + \sigma W_t}\right] = S_0 e^{\left(\mu - \frac{\sigma^2}{2}\right)t} E\left[e^{\sigma W_t}\right]$$
$$\Leftrightarrow E[S_t] = S_0 e^{\mu t}, \tag{2.10}$$

i.e. the asset price is expected to grow with a continuously compounded interest rate.[6]

2.3 Risk-neutral valuation of contingent claims

We now define a two-security world where the first security is a riskless bond, whose price evolves at a constant *risk-free interest rate*[7] r from its initial value $B_0 = 1$ according to the differential equation

$$dB_t = rB_t dt. \tag{2.11}$$

Its price at time t is thus equal to the continuously compounded risk-free interest rate: $B_t = e^{rt}$. The second security is a risky asset, whose price evolves from its initial value S_0 following an Itō process

$$dS_t = \mu S_t dt + \sigma S_t dW^{\mathbb{P}}, \tag{2.12}$$

where μ and σ are drift and volatility. $W^{\mathbb{P}}$ is a Wiener process under the probability measure $\mathbb{P}$ for which $\mathbb{P}(\omega) > 0 \ \forall \omega \in \ \Omega$ and $\sum_{\omega \in \Omega} \mathbb{P}(\omega) = 1$, where Ω is a nonempty space denoting all possible states ω of the world: $\Omega = \{\omega_1, \omega_2, \dots\}$. By Itō's lemma and equations (2.11) and (2.12), the discounted asset price process $Z_t = S_t/B_t = S_t e^{-rt}$ is shown to equal

$$dZ_t = (\mu - r)Z_t dt + \sigma Z_t dW^{\mathbb{P}}. \tag{2.13}$$

Define the drift process of the above equation as $(\mu - r)Z_t = \sigma\lambda Z_t$, so that $\lambda = (\mu - r)/\sigma$. Rearranging shows that the unknown drift parameter of the risky asset's price process is equal to the risk-free interest rate plus its volatility weighted by λ. This weighting factor

[6] Continuous compounding is the limit when the assumed compounding interval becomes infinitesimal. In the context above, if μ is assumed to equal the interest rate *per annum* and t measures time in years, the future value of the asset price would be equal to $S_0(1 + \frac{\mu}{n})^{nt}$ when compounding takes place n times per annum. Continuous compounding is then given by $\lim_{n\to\infty} S_0 \left(1 + \frac{\mu}{n}\right)^{nt} = S_0 e^{\mu t}$.

[7] In theory, the risk-free rate is the price of riskless money, i.e. the rate at which money can be borrowed (or lent) without credit risk, so that it is certain to be repaid. The Treasury rate (the rate at which a government borrows in its own currency) is often thought as risk-free. Practitioners, however, usually set the risk-free rate equal to an Interbank Offer Rate, such as the Libor or the Euribor. (Hull 2009.)

then represents the *market price of risk* of the risky asset S_t: the excess return required as compensation for an additional unit of standard deviation. Clearly, for a riskless asset such as bond B_t we have $\lambda = 0$ and thus $\mu = r$. By applying the Girsanov theorem[8], we can define a new probability measure under which the market price of risk is always zero and all asset prices thus grow at a constant (risk-free) interest rate. This implies that investors' risk preferences have no impact on their investment decisions; under this new probability measure, the world is *risk-neutral*.

To this end, let $\tilde{W}^{\mathbb{Q}}$ be defined by $d\tilde{W}^{\mathbb{Q}} = \lambda dt + dW^{\mathbb{P}}$. The Girsanov theorem then states that $\mathbb{Q}$ is an equivalent probability measure such that $\tilde{W}^{\mathbb{Q}}$ is a Wiener process. Under this new measure, the discounted price return process (2.13) can be represented as a zero drift process and the asset price process has a drift parameter equal to the risk-free interest rate:

$$\begin{aligned} &\frac{dZ_t}{Z_t} = \sigma\lambda dt + \sigma dW^{\mathbb{P}} = \sigma d\tilde{W}^{\mathbb{Q}} \\ \Leftrightarrow\ &\frac{dS_t}{S_t} = r dt + \sigma d\tilde{W}^{\mathbb{Q}}. \end{aligned} \tag{2.14}$$

From the above formulation it is easy to see that in a risk-neutral world (under the probability measure $\mathbb{Q}$) the volatility parameter represents the only source of uncertainty in the asset price process.

The fundamental theorem of asset pricing (see e.g. Harrison and Pliska 1981, Delbaen and Schachermayer 1994) states that the existence of $\mathbb{Q}$ implies the absence of arbitrage opportunities. An arbitrage is defined as the opportunity to make a risk-free profit at zero cost, i.e. there is no probability of loss, the probability of gain is positive and no net investment of capital is required. Formally, an arbitrage opportunity is said to exist if the value of a portfolio Π at time t satisfies $\Pi_0 = 0$, $\mathbb{P}(\Pi_t \geq 0) = 1$ and $\mathbb{P}(\Pi_t > 0) > 0$.

The absence of arbitrage implies that all assets with identical future cash flows have an identical price.[9] Therefore, the future values of all asset prices are defined by the

[8] The theorem attributed to Girsanov (1960) allows to express the dynamics of a stochastic process under a new, equivalent probability measure denoted here by $\mathbb{Q}$. Equivalency requires that $\mathbb{P}(\omega) > 0 \Leftrightarrow \mathbb{Q}(\omega) > 0 \quad \forall\omega \in \Omega$. Essentially, the Girsanov theorem links the probability measures using a Radon-Nikodym derivative (see Nikodym 1930):

$$E^{\mathbb{Q}}[X_t] = E^{\mathbb{P}}\left[\frac{d\mathbb{Q}}{d\mathbb{P}} X_t\right],$$

where X_t is a Wiener process.

[9] This can be seen by considering a situation where arbitrage opportunities existed due to e.g. discrepancies between different markets: An investor willing to benefit from the arbitrage could buy the cheaper

continuously compounded risk-free interest rate (recall equation (2.10)): $E^{\mathbb{Q}}[S_t] = S_0 e^{rt}$. For the return of the risky asset S_t over a time period $\tau = T - t$, this translates to

$$E^{\mathbb{Q}}\left[\frac{S_T}{S_t}\Big|S_t\right] = e^{r(T-t)}, \tag{2.15}$$

as was first noted by Cox and Ross (1976).

The theory of rational option pricing (Merton 1973) states that the price of a derivative must equal its risk-neutral expected payoff at maturity. A *forward price* F_t can thus be deduced directly from above and is given by

$$F_t = E^{\mathbb{Q}}[S_T] = S_t e^{r(T-t)}. \tag{2.16}$$

Arbitrage considerations show that if $F_t > S_t e^{r(T-t)}$, profit could be made by buying the asset S and shorting the forward contract F on that asset. Conversely, if $F_t < S_t e^{r(T-t)}$, an opposite strategy would be profitable. Therefore, under the assumption of no arbitrage, the equality must hold.

For option valuation using risk-neutral pricing, the expected payoffs at maturity are discounted by the continuously compounded risk-free interest rate. At any time t before the expiry date T the theoretical European call and put option prices $C_t(K,T)$ and $P_t(K,T)$ written on asset S are equal to these discounted payoffs, i.e.

$$\begin{aligned} C_t(T,K) &= e^{-r(T-t)} E^{\mathbb{Q}} (S_T - K)^+ \\ &= e^{-r(T-t)} \int_0^\infty (S_T - K)^+ f(S_T, T|S_t, t) dS_T \end{aligned} \tag{2.17}$$

and

$$\begin{aligned} P_t(T,K) &= e^{-r(T-t)} E^{\mathbb{Q}} (K - S_T)^+ \\ &= e^{-r(T-t)} \int_0^\infty (K - S_T)^+ f(S_T, T|S_t, t) dS_T, \end{aligned} \tag{2.18}$$

where $f(S_T, T|S_t, t)$ is the conditional probability distribution function or the risk-neutral density (RND) of the underlying asset price at expiry, S_T, given its price at time t. Equations (2.17) and (2.18), together with equation (2.15), provide the solution to the option valuation problem (Ross 1976, Cox and Ross 1976; see also Harrison and Pliska 1981).

asset (with identical future cash flows) and sell it on another market at a higher price. But other investors would then quickly notice and take advantage of these opportunities, which would lead to them disappearing. Therefore, in the long run it is reasonable to assume that an equilibrium with a single price for all assets would arise.

The RND is unknown and needs to be estimated. An extensive array of parametric estimation methods exist for this purpose, probably the most famous among them being the model by Black and Scholes (1973) (see also Merton 1973), who give an analytical solution for European option prices under certain restrictive and perhaps somewhat simplistic assumptions on the asset price process and market environment. The power of the model, however, lies in its robustness and ease-of-use: All parameters save for one are readily available from market quote providers, and the formula normally gives a theoretical price close enough to the actual value of the option. The original article of Black and Scholes (1973) sparked a boom of interest in option pricing problems, and several extensions and alternative estimation methods have been developed in the succeeding years. For up-to-date guides to option pricing models and their implementation, see e.g. Haug (2007) and Rouah and Vainberg (2007).

The following chapter discusses the concept of financial volatility. As a measure of the uncertainty of the realized price return of an asset (recall that it represents the only source of risk in an asset return process under the risk-neutral measure), it is an integral component of option pricing. Conversely, volatility can be obtained from (observed) option prices since it is embedded in them, which dismisses the need for an option pricing model.

Chapter 3

Gauging volatility

Financial market volatility refers to the variability or dispersion of price changes over time. It is not therefore a measure of how much prices are changing; an asset whose price is consistently increasing or decreasing by a large but similar amount each period would have low price volatility. Instead, volatility measures how price movements themselves vary over time – it is a gauge of the uncertainty of an asset's returns. The higher the volatility, the riskier the security, since heightened volatility translates to a larger spectrum of potential prices. This also increases the value of options written on the asset.

When defining volatility, an important distinction is between *historic* and *implied* volatility. Historic volatility is a statistical measure of the dispersion of asset returns. It is backward-looking by definition, since it is calculated from past (historical) data, usually a set of daily observations drawn from a period such as one calendar month or, 22 trading days. As a statistical measure, volatility corresponds to standard deviation and is consequently commonly denoted by σ. A consistent estimator of σ is the sample standard deviation of asset returns (see e.g. Hull 2009). It has prevailed as the most common measure for historical volatility largely because it is consistent regardless of the data distribution, i.e. it converges to σ as the number of observations tends to infinity.[1]

In contrast, implied volatility can be perceived as the value of volatility that, when

[1] An alternative volatility measure is given by the mean absolute deviation (MAD), which in its unadjusted form is biased and needs to be scaled according to the observed data distribution, rendering it considerably more inconvenient. However, Ederington and Guan (2006) found their adjusted MAD to be a better volatility estimate than the historical standard deviation when working with normally distributed data. Goldstein and Taleb (2007), for their part, argue that finance professionals often confuse mean absolute deviation with standard deviation when talking about volatility, which leads to a significant underestimation of the actual (realized) volatility.

inserted into an option pricing formula, equates the theoretical price of the option to its market price. It is the level of volatility that is implied by the current market price of the option together with all other determinants affecting the value of an option in an option pricing framework. Therefore, assuming that the markets process information efficiently, implied volatility can be thought of as the expected level of volatility over the remaining life of the option (its time to expiry). As opposed to historical volatility, it is inherently a forward-looking measure and provides a consensus forecast of the expected level of volatility in a future time period.

Implied volatility is, on average, higher than historic volatility (see e.g. Carr and Wu 2006; 2009). This can be explained by the volatility risk premium, which is essentially the compensation paid to option sellers for bearing the risk of losses during periods in which actual volatility increases suddenly, without being incorporated into implied volatility through option prices. Implied volatility is generally found to be negatively correlated with asset returns, even though volatility as a measure does not reveal anything of the direction of price changes. The relationship between volatility and asset returns is delved into more deeply in the next chapter.

Recent research has showed that implied volatility can be derived from observed market option prices independently of any option pricing model. This is favorable since economic models are by definition simplified representations of complex processes and often posit structural parameters, they are subject to misspecification errors. Tests based on a model-free measure of implied volatility are direct tests of market efficiency rather than joint tests of market efficiency and the assumed option pricing model (Jiang and Tian 2005). In the next section we examine the concept of model-free implied volatility more closely.

3.1 Model-free implied volatility

Unlike the traditional concept of implied volatility, where the implied volatility is estimated numerically from an option pricing model, the model-free implied volatility (MFIV) does not rely on any particular parametric model – instead, it can be calculated directly from the market prices of a cross section of call and put options with the same expiry. The concept was originally proposed by Dupire (1993) and Neuberger (1994)[2]

[2] See also Carr and Madan (1998) and Demeterfi et al. (1999).

and was comprehensively formulated first in Britten-Jones and Neuberger (2000). The authors demonstrate that the set of call and put option prices (defined exogenously in the markets) having the same maturity is sufficient to derive the risk-neutral expected sum of the squared returns of the asset between the current date and the option maturity. In their original derivation they assume zero interest rates and consider an asset that pays no dividend. They make no further assumptions regarding the underlying stochastic process of the asset except the fact that both the asset and the volatility exhibit no jumps. Jiang and Tian (2005) further refine the original result by showing how to implement the concept of model-free implied volatility in the more real-world-like setting of dividends and non-zero interest rates. They also prove that the restriction of no jumps in the asset prices can be relaxed as well. In their derivation of a closely related measure referred to as the corridor implied volatility, Andersen and Bondarenko (2007) allow for jump discontinuities in the volatility process by defining it as a càdlàg function[3].

Britten-Jones and Neuberger (2000) derive their result based on earlier results proposed in Ross (1976), Banz and Miller (1978) and Breeden and Litzenberger (1978), who discovered that the RND can be recovered from a set of European option prices. The relationship suggests that the RND is simply the second partial derivative of the call (or put) pricing function with respect to the strike price, and it allows expressing the risk-neutral expected sum of squared returns via the RND using only observed option prices and thus without the need to specify a pricing function for the options. In their original derivation, Britten-Jones and Neuberger (2000) assume that the price return process is fully defined by the instantaneous volatility:

$$\frac{dS_t}{S_t} = \sigma(t, \cdot)dz. \tag{3.1}$$

The risk-neutral expected sum of squared returns between dates t_1 and t_2 is then[4]

$$E\left[\int_{t_1}^{t_2} \left(\frac{dS_t}{S_t}\right)^2 dt\right] = 2\int_0^{\infty} \frac{C(t_2, K) - C(t_1, K)}{K^2} dK. \tag{3.2}$$

The subscript $\mathbb{Q}$ is omitted from the expectation operator above for simplicity. Throughout the rest of this thesis, the expectation operator will refer specifically to risk-neutral expectation, unless stated otherwise. A model-free measure of volatility is then given as the square root of the above expression. Notice that this formulation makes use of the en-

[3]"Continue à droite, limite à gauche", meaning "continuous on (the) right, limit on (the) left".
[4]For details on the derivation, see Britten-Jones and Neuberger (2000, appendix A).

tire range of strike prices (theoretically from zero to infinity), and therefore presumably captures all information within the options of a given expiry.

Jiang and Tian (2005) derive an expression of the MFIV that is valid for a more general class of asset price processes (e.g. ones that contain jumps). Dividends can be accounted for by considering an asset price $S_t^* = S_t - PV_t(D)$, where the present value PV_t of total dividends paid D is removed from the initial asset price. To relax the assumption of zero interest rates, the authors employ a forward asset $F_t = S_t^* e^{rT}$, which has a zero drift rate in a risk-neutral world. Setting $t_1 = 0$ and $t_2 = T$ in Equation (3.2) above yields the realized asset return variance between now and some future expiry date. As Jiang and Tian (2005) demonstrate, the risk-neutral expected sum of squared returns for a forward asset F_t is then given by

$$E\left[\int_0^T \left(\frac{dF_t}{F_t}\right)^2 dt\right] = 2\int_0^\infty \frac{C^F(T, Ke^{rT}) - (F_0 - K)^+}{K^2} dK, \tag{3.3}$$

where $C^F(T, K) = e^{rT} C(T, K)$, the forward option price, and $(F_0 - K)^+ = C^F(0, K)$, the time-zero value of the forward call option. By redefining $S_t = S_t^*$, Equation (3.2) can then be stated as

$$E\left[\int_0^T \left(\frac{dS_t}{S_t}\right)^2 dt\right] = 2\int_0^\infty \frac{e^{rT} C(T, K) - (S_0 e^{rT} - K)^+}{K^2} dK. \tag{3.4}$$

As shown in Appendix A.1, this can further be modified to yield a more compact expression for the MFIV:

$$\sigma^2_{\text{MFIV}} = 2e^{rT}\left[\int_0^{F_0} \frac{P(T, K)}{K^2} dK + \int_{F_0}^\infty \frac{C(T, K)}{K^2} dK\right]. \tag{3.5}$$

The MFIV can thus theoretically be computed by taking a cross section of call and put option prices over an infinite range of strike prices. In practice, the computation requires some approximation owing to the discrete nature of financial markets and the finite range of observable market option prices for any given security. The VIX volatility index introduced in Section 3.3 is a practical implementation of MFIV, although its theoretical foundations lay in the pricing of variance swaps, which is discussed in the next section.

3.2 Variance swap as model-free implied volatility

The VIX volatility index introduced in Section 3.3 is based on the theory of variance swaps. Therefore, in spite of its similarities with the concept of model-free implied volatility, the theoretical pricing of variance swaps merits a closer look. A variance swap is a forward contract on the future realized variance of an underlying asset. At expiry, it has a payoff equal to $\sigma_R^2 - K_{\text{Var}}$ times the notional value of the swap. σ_R^2 is the realized asset variance quoted in annual terms over the life of the contract and K_{Var} is the strike price. Demeterfi et al. (1999) define the fair delivery value of future realized variance (henceforth referred to as FVFV) as the strike price that is equal to the risk-neutral expected value of average future variance. The concept of the FVFV is closely related to the MFIV – alas, their theoretical equivalence was is proven by Jiang and Tian (2007). Its derivation is based on a log contract (see Neuberger 1994), which can be perfectly replicated and hence priced using a portfolio of call and put options. This makes it easier to implement in practice.

Following Demeterfi et al. (1999) (see also Bossu, Strasser, and Guichard 2005), we begin by assuming that the evolution of the underlying price S_t is given by the following Itō process:

$$dS_t = \mu_t S_t dt + \sigma_t S_t dW_t \Leftrightarrow \frac{dS_t}{S_t} = \mu_t dt + \sigma_t dW_t \tag{3.6}$$

with time-varying mean and volatility parameters μ_t and σ_t. Notice that the underlying asset's price evolution given above is a more general version of the geometric Brownian motion discussed in Section 2.2.

By taking logarithms and applying Itō's lemma with $f(S,t) = \ln S_t$ to (3.6) we obtain

$$\begin{aligned} d(\ln S_t) &= \left(S_t \mu_t \frac{1}{S_t} - \frac{1}{2} S_t^2 \sigma_t^2 \frac{1}{S_t^2}\right) dt + S_t \sigma_t \frac{1}{S_t} dW_t \\ &= \left(\mu_t - \frac{1}{2}\sigma_t^2\right) dt + \sigma_t dW_t && (3.7) \\ \Leftrightarrow \frac{1}{2}\sigma_t^2 dt &= \frac{dS_t}{S_t} - d(\ln S_t) && (3.8) \end{aligned}$$

where σ_t^2 is the (instantaneous) variance at time t. Setting the current time at $t = 0$, the average variance over a continuous sample between now and a future time T is then given by the continuous integral

$$\mathrm{Var}_T = \frac{1}{T}\int_0^T \sigma_t^2 dt. \tag{3.9}$$

From Equation (3.7),

$$\mathrm{Var}_T = \frac{2}{T}\left(\int_0^T \frac{dS_t}{S_t} - \int_0^T d(\ln S_t)\right) = \frac{2}{T}\left[\int_0^T \frac{dS_t}{S_t} - \ln\left(\frac{S_T}{S_0}\right)\right]. \tag{3.10}$$

The FVFV is defined as the expected value of the average variance under the risk-neutral measure: $K_{\mathrm{Var}} = E[\mathrm{Var}_T]$. Recall that in a risk-neutral world the drift rate μ_t in equation (3.6) equals a constant risk-free interest rate r, so that the rate of return of the underlying asset equals the risk-neutral expected return over $[0,T]$, viz. rT. Thus, we obtain

$$\begin{aligned} K_{\mathrm{Var}} = E[\mathrm{Var}_T] &= E\left[\frac{1}{T}\int_0^T \sigma_t^2 dt\right] \\ &= \frac{2}{T}E\left[\int_0^T \frac{dS_t}{S_t} - \ln\left(\frac{S_T}{S_0}\right)\right] \\ &= \frac{2}{T}\left(rT - E\left[\ln\left(\frac{S_T}{S_0}\right)\right]\right) \end{aligned} \tag{3.11}$$

The final term in the last identity, $E\left[\ln\left(\frac{S_T}{S_0}\right)\right]$, represents the expectation of the payoff function of a log contract, which can be replicated using a forward contract and a set of put and call options, all with a common expiration date.[5] To this end, the log-payoff function is first decomposed using an arbitrary parameter S_* defining the boundary, or cutoff point, between calls and puts:

$$\ln\left(\frac{S_T}{S_0}\right) = \ln\left(\frac{S_T}{S_*}\right) + \ln\left(\frac{S_*}{S_0}\right) \tag{3.13}$$

where the second term on the right-hand-side (RHS) is a constant. Therefore, only the

[5] The replication argument derived by Carr and Madan (1998) states that any twice continuously differentiable function $f(S)$ defined in $\mathbb{R}$ can be replicated as

$$f(S) = f(\kappa) + f'(\kappa)(S-\kappa) + \int_0^\kappa f''(K)(K-S)^+ dK + \int_\kappa^\infty f''(K)(S-K)^+ dK \tag{3.12}$$

for some threshold κ and strike price K (see Appendix A.2 for proof).

first term on the RHS has to be replicated. The replication is given by

$$\ln\left(\frac{S_T}{S_*}\right) = \frac{S_T - S_*}{S_*} - \int_0^{S_*} \frac{(K - S_T)^+}{K^2} dK - \int_{S_*}^{\infty} \frac{(S_T - K)^+}{K^2} dK. \tag{3.14}$$

This formulation can be interpreted as holding a portfolio consisting of a short position in $1/S_*$ forward contracts with strike S_* and short positions in $1/K^2$ put options for all strikes $K \in [0, S_*]$ and another $1/K^2$ in call options for all strikes $K \in [S_*, \infty[$. Recalling the theoretical prices of forward and option contracts defined in Section 2.3, the final term of equation (3.11) can then be stated as

$$\begin{aligned}
-E\left[\ln\left(\frac{S_T}{S_0}\right)\right] &= -\ln\left(\frac{S_*}{S_0}\right) - E\left[\ln\left(\frac{S_T}{S_*}\right)\right] \\
&= -\ln\left(\frac{S_*}{S_0}\right) - \frac{E[S_T] - S_*}{S_*} + \int_0^{S_*} \frac{E[(K - S_T)^+]}{K^2} dK + \int_{S_*}^{\infty} \frac{E[(S_T - K)^+]}{K^2} dK \\
&= -\ln\left(\frac{S_*}{S_0}\right) - \left(\frac{S_0}{S_*} e^{rT} - 1\right) \\
&\quad + e^{rT} \left(\int_0^{S_*} \frac{P(T,K)}{K^2} dK + \int_{S_*}^{\infty} \frac{C(T,K)}{K^2} dK\right)
\end{aligned} \tag{3.15}$$

Substituting the above expression for $-E\left[\ln\left(\frac{S_T}{S_0}\right)\right]$ in equation (3.11) gives the fair value of future variance:

$$\begin{aligned}
K_{\text{Var}} = \frac{2}{T} \Bigg[& rT - \left(\frac{S_0}{S_*} e^{rT} - 1\right) - \ln\left(\frac{S_*}{S_0}\right) \\
& + e^{rT} \left(\int_0^{S_*} \frac{P(T,K)}{K^2} dK + \int_{S_*}^{\infty} \frac{C(T,K)}{K^2} dK\right) \Bigg].
\end{aligned} \tag{3.16}$$

This is Equation (26) of Demeterfi et al. (1999). As shown by Jiang and Tian (2007), the theoretical equivalence of this formulation with the concept of model-free implied is obtained by setting the integral cutoff point to equal to the forward price, i.e. $S_* = F_T = S_0 e^{rT}$ (see Appendix A.1 for details on the calculation).

The concepts of MFIV and FVFV derived above make precise the perhaps intuitive notion that implied volatilities can be regarded as the market's expectation of future realized volatilities. They establish a connection between market prices of options and the risk-neutral expected realized variance (squared volatility), but their practical implementation requires some approximations, as can be seen from e.g. the continuous integrals on strike prices in Equation (3.16). In the next section we examine how the FVFV is approximated and implemented in practice in the VIX volatility index.

3.3 The VIX volatility index

This section introduces the current computation methodology of the VIX volatility index first implemented by the Chicago Board Options Exchange (CBOE) in 1993. A volatility index for measuring the aggregate level of stock market volatility was first proposed by Gastineau (1977), shortly after the first option contracts began trading on the CBOE in April 1973. Gastineau (1977) used an average of at-the-money options on 14 stocks combined with a measure of historical volatility, and Cox and Rubinstein (1985) refined this procedure by including multiple calls on each stock and by weighting the volatilities in such a way that the index is always at-the-money and has a constant time to expiration. (Fleming, Ostdiek, and Whaley 1995). Brenner and Galai (1989) constructed volatility indices for equity, bond and foreign exchange markets and Whaley (1993) introduced the computation methodology that was later implemented by the CBOE as the first exchange to disseminate an official volatility index on a regular basis. The success of this index sparked interest in other exchanges around the world, who then developed their own respective indices. Effectively all introduced indices tracking market volatility on different markets make use of the original methodologies.[6]

Initially, the VIX was based on the implied volatilities of the S&P 100 stock index, which at the time possessed the most actively traded stocks on the US market and was therefore considered its foremost benchmark. After the introduction of the new VIX, the CBOE has continued disseminating the original volatility index under a new ticker symbol, VXO. This first implementation is based on the implied volatilities of four pairs of near-the-money[7] call and put options of the nearest and second-nearest maturity on the S&P 100 stock index options. The implied volatilities are derived from the popular Black and Scholes (1973)/Merton (1973) option valuation model. Since these options are American and since the underlying OEX index portfolio pays discrete cash dividends,

[6]For reviews on volatility indices in different markets, see Aboura and Villa (1999) and Siriopoulos and Fassas (2009).

[7]From the payoff functions of (European) call and put options it is clear that the owner has the possibility to accrue unlimited profit at a limited risk of loss, and the option has intrinsic value relative to its *moneyness*. An option is said to be *in-the-money* (ITM) if it has some intrinsic value, that is, if $K < S_t$ for a call and $K > S_t$ for a put. An option is *out-of-the-money* (OTM) if it has no intrinsic value, and *at-the-money* (ATM), if $K = S_t$. The somewhat vaguer term *near-the-money* means simply that the strike price is close to the underlying's current market price.

the implied volatilities need to be estimated numerically. The method of choice is the binomial method of Cox, Ross, and Rubinstein (1979) adjusted for cash dividends, as described in Harvey and Whaley (1992). The implied volatilities are weighted in such a manner that the volatility index represents a measure of the market consensus of the expected volatility over the next 30 *calendar days*. Moreover, since the implied volatilities used in the index calculation are stated in *trading days*[8], an artificial trading-day adjustment is introduced to correctly calibrate the index value (for details on the calculation, see Whaley (1993; 2000) and Fleming et al. (1995)).

The trading day adjustment of the original computation methodology leads to an upward bias in the volatility index value. It has drawn criticism from both academia and industry and was one of the main reasons behind the CBOE's decision in 2003 to introduce a new volatility index with revamped computation methodology (Carr and Wu 2006). The new VIX is based on the work of Demeterfi et al. (1999) and is closely related to the concept of model-free implied volatility. It approximates the 30-day variance swap rate and accommodates a broader range of options, thus capturing information from the entire volatility smile rather than the implied volatilities of only at-the-money options (CBOE 2009). Moreover, the index is computed directly from observable option prices and therefore does not rely on any specific option model. The calculation methodology allows for replicating implied volatility with a static portfolio of options, which has led to the creation of tradable products on the VIX. Lastly, the new VIX uses options data on the S&P 500 index, the current core index of US equities that is considered the best representation of the market as a whole. Historical data for the original index is available from 1986 to the present, whereas the new VIX has been backdated to 1990 to facilitate comparison between the two indices.

In the new specification introduced in 2003 the index values are calculated directly from market prices of the underlying index options instead of using Black-Scholes implied volatilities. This new calculation method is consistent with the theory of model-free implied volatility, although the methodology relies on the concept of fair value of future variance. More precisely, the VIX is calculated as a discretized approximation of the fair strike of a variance swap (K_{Var} in Section 3.2 above) with a fixed 30-day maturity on the S&P 500 stock index.

In the VIX calculation procedure CBOE uses several approximations in order for the

[8]There are approximately 252 trading days in a year.

VIX to correctly represent the market's volatility expectation over the next 30 calendar days. These approximations are given below following Jiang and Tian (2007). We begin with the formulation of FVFV derived in Section 3.2 and restated here for convenience.

$$\begin{aligned} K_{\mathrm{Var}} = \frac{2}{T} \Bigg[& rT - \left(\frac{S_0}{S_*} e^{rT} - 1 \right) - \ln \left(\frac{S_*}{S_0} \right) \\ & + e^{rT} \left(\int_0^{S_*} \frac{P(T,K)}{K^2} dK + \int_{S_*}^{\infty} \frac{C(T,K)}{K^2} dK \right) \Bigg] . \end{aligned} \tag{3.17}$$

First, as can be seen from the integrals in the expression, the variance calculation requires an infinite range of strike prices for the put and call options. However, in financial markets there is, of course, only a finite range of strike prices available for trading on any given underlying with any given maturity. Therefore, the infinite range of strike prices has to be replaced with a range from the lowest available strike price, K_L, to the highest available strike K_H for a given maturity.

The integral cutoff point is set at $S_* = K_0$, where K_0 is the first available price below the forward index level F_0 defined in the usual way by equation (2.16).[9] This leads to the following approximation:

$$\int_0^{S_*} \frac{P(T,K)}{K^2} dK + \int_{S_*}^{\infty} \frac{C(T,K)}{K^2} dK \approx \int_{K_L}^{K_0} \frac{P(T,K)}{K^2} dK + \int_{K_0}^{K_H} \frac{C(T,K)}{K^2} dK. \tag{3.19}$$

To account for the discrete nature of the financial markets, the continuous integrals need to be replaced with a discrete approximation using option prices quoted for trading (there is only a finite range of strike prices at given increments to choose from between K_L and K_H). The numerical method used by the CBOE to approximate the continuous integrals is described as

$$\int_{K_L}^{K_0} \frac{P(T,K)}{K^2} dK + \int_{K_0}^{K_H} \frac{C(T,K)}{K^2} dK \approx \sum_i \frac{\Delta K_i}{K_i^2} Q(T, K_i), \tag{3.20}$$

where $Q(T, K_i)$ is the midpoint of bid and ask price quotes for the option and ΔK_i is the

[9]In practice, the CBOE determines the forward index level by choosing call and put options with prices that are closest to each other. The forward price is then derived via a convenient relation known as the put-call parity: $C(K,T) - P(K,T) = S_0 - Ke^{-rT}$. Therefore,

$$F_0 = S_0 e^{rT} = e^{rT}(C(T,K) - P(T,K)) + K. \tag{3.18}$$

(Carr and Wu 2006).

strike price increment given by

$$\Delta K_i = \frac{K_{i+1} - K_{i-1}}{2}$$

and K_i is the strike price of the i-th option (a call if $K_i > F_0$, a put otherwise). This price increment is modified at K_H and K_L and redefined as the difference between the two highest and the two lowest strike prices, respectively. Another adjustment is made at strike K_0, where $Q(T, K_i)$ is redefined as the average price of a call and a put.

The final approximation is due to a Taylor series expansion[10] the CBOE uses in its calculation procedure to estimate the log function in (3.17) in terms of the forward index level F_0 and strike prices. Notice that the terms preceding the sum of integrals in (3.17) can be expressed as

$$\begin{aligned} \frac{2}{T}\left[rT - \left(\frac{S_0}{S_*}e^{rT} - 1\right) - \ln\frac{S_*}{S_0}\right] &= \frac{2}{T}\left[rT - \frac{F_0}{K_0} - 1 - \ln K_0 + \ln\left(\frac{F_0}{e^{rT}}\right)\right] \\ &= \frac{2}{T}\left[\ln\left(\frac{F_0}{K_0}\right) - \left(\frac{F_0}{K_0} - 1\right)\right]. \end{aligned} \tag{3.21}$$

Applying the Taylor series expansion on the log function and ignoring all but the first and second order terms gives the following approximation

$$\ln\left(\frac{F_0}{K_0}\right) \approx \left(\frac{F_0}{K_0} - 1\right) - \frac{1}{2}\left(\frac{F_0}{K_0} - 1\right)^2 \tag{3.22}$$

Substituting the RHS of the above approximation for $\ln\left(\frac{F_0}{K_0}\right)$ in the second row of equation (3.21) yields

$$\frac{2}{T}\left[rT - \left(\frac{S_0}{S_*}e^{rT} - 1\right) - \ln\left(\frac{S_*}{S_0}\right)\right] \approx -\frac{1}{T}\left(\frac{F_0}{K_0} - 1\right)^2. \tag{3.23}$$

Andersen, Bondarenko, and Gonzalez-Perez (2011) point out that this final approximation acts as a (negative) correction term for the discrete approximation. The final form of the CBOE model-free implied variance measure, from which the VIX index value is derived, is obtained by substituting the approximations (3.19), (3.20) and (3.23) into the

[10] A one-dimensional Taylor series expansion of a real function $f(x)$ around a point $x = a$ is given by

$$f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \frac{f^{(3)}(a)}{3!}(x-a)^3 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n + \ldots$$

FVFV in equation (3.17). Therefore, as explained in CBOE (2009), the formula for calculating the VIX index value is given by

$$\hat{\sigma}^2_{\mathrm{VIX}} = \frac{2}{T}\sum_i \frac{\Delta K_i}{K_i^2} e^{rT} Q(K_i) - \frac{1}{T}\left(\frac{F_0}{K_0} - 1\right)^2, \tag{3.24}$$

where $\hat{\sigma}^2_{\mathrm{VIX}}$ is the VIX variance measure. The VIX index itself is defined for a fixed calendar maturity of $T_M = \frac{30}{365}$, or thirty days – however, at any given time there are generally no options that expire in exactly 30 days. The CBOE's solution is linear interpolation from two maturities, T_1 and T_2, that are closest to the required 30 days:

$$\hat{\sigma}^2_{\mathrm{VIX}}(T_M) = \frac{1}{T_M}\left[\omega T_1 \sigma^2_{vix}(T_1) + (1-\omega) T_2 \sigma^2_{vix}(T_2)\right], \quad \omega = \frac{T_2 - T_M}{T_2 - T_1}. \tag{3.25}$$

Options with less than seven days to expiry are excluded to minimize pricing anomalies that might occur close to expiration.[11] Finally, the VIX index value is obtained as the annualized quantity of the above interpolation, multiplied by 100 to yield a quote in volatility units (percentage points):

$$\mathrm{VIX} = 100\sqrt{\frac{30}{365}\hat{\sigma}^2_{\mathrm{VIX}}(T_M)}. \tag{3.26}$$

Jiang and Tian (2005) find in their study of the information content of model-free implied volatility that it subsumes all information contained in historical volatility as well as in the Black-Scholes implied volatility on which the VIX was previously based. This renders the new model-free VIX a more efficient forecast for future volatility at least in theory and lends further support to the new calculation methodology. However, Jiang and Tian (2007) assess the nature and significance of the measurement errors caused by the overall approximation scheme in the CBOE procedure decomposed above. They classify the errors as (i) truncation errors – the minimum and maximum available strike prices, K_L and K_H, are far from zero and infinity required by theory; (ii) discretization errors – the continuous integrals are approximated by piecewise linear functions; (iii) expansion errors – a log function is approximated using only the first and second order terms of a Taylor series expansion; and (iv) interpolation errors arising from the fixed 30-day maturity requirement. The authors find that these errors can lead to substantial and economically significant biases in the calculated VIX index value (see also Carr and Wu 2009)

[11] See the white paper by CBOE (2009) for details and practical examples on the calculation.

and propose a simple "smoothing method" to overcome the implemention issues. Any corrections to the VIX methodology are yet to be implemented by the CBOE.

This chapter concludes our discussion of the theories and methodology that constitutes the VIX volatility index in its current specification. In the remaining chapters we examine the empirical properties of the VIX index time series, the regime shifts therein as well as the relationship between implied volatility and stock index returns.

Chapter 4

Volatility regimes: Identifying structural breaks in the VIX time series

The currently used, revised calculation methodology for the VIX was introduced in September 22, 2003. Upon introducing the updated methodology, CBOE created a historical record for the new VIX dating back to 1990 to facilitate comparison with its predecessor now denoted the VXO. In this thesis, however, we consider only the quotes that have been disseminated in real time and hence our daily data series, obtained from Yahoo! Finance, extends from January 2, 2004 to December 31, 2011 and consists of the daily closing levels for 2015 trading days. The choice of sample period also coincides with the introduction of the first exchange-traded derivative instruments on volatility (VIX futures in early 2004; VIX options were subsequently launched in 2006), for which the new calculation methodology provided a basis. The sample period ranges from the years of relative calm (pre-2007) to the recent times of market turbulence covering the outbreak of the financial crisis in 2007-2008 and the subsequent global recession.

The time series sample is graphed in Figure 4.1. Clearly, the VIX movement was encapsuled in a relatively tight range until 2007 and the outbreak of the global financial crisis. Thereafter, and especially from 2008 onwards, the VIX seems to have shifted to a regime with higher mean level and more pronounced movements, and with notable spikes to the upside. Visual inspection of the time series suggests structural breaks in the data – we are interested in ascertaining whether there are distinctive regimes in volatility, e.g. a low-volatility state or a high-volatility state characterized by calm or even complacency and financial stress, respectively. An antecedent is given by Guo and Wohar (2006), who studied the VIX and its predecessor, the VXO, for the time periods 1990-2003 and 1986-2003, respectively. They find evidence of three distinctive regimes in both series with the breaks occurring in circa 1992 and 1997. This chapter thus provides an update using more recent data spanning periods of abnormal levels of volatility evident in Figure 4.1. Identifying distinct regimes in volatility allows us to gain insight on volatility dynamics. What's more, it permits the study of the relationship between stock market volatility and price returns in differing market environments.

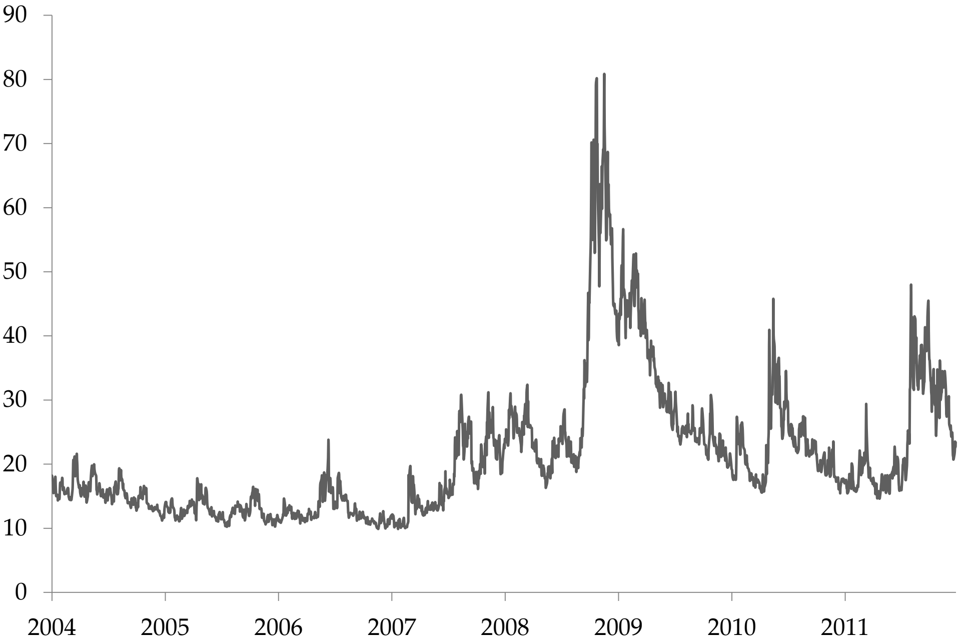

Figure 4.1: VIX daily time series from January 2004 to December 2011.

4.1 Normal ranges

Before examining the regime changes in volatility in a more rigorous manner, we attempt to characterize the behavior of the stock market volatility in a probabilistic sense. Following the approach of Whaley (2000; 2009), the normal ranges of the VIX are given in Table 4.1. The percentiles represent the boundaries below which the given percentages of observations fall. For example, the 50th percentile separates the lower half of observations from the higher half, i.e. it is the sample median within each time interval, while the 5th and 95th percentiles mark the boundaries for the most extreme values observed.

Table 4.1: Yearly percentile ranges for VIX daily closing levels from January 2, 2004 to December 31, 2011.

		Percentiles						
Year	# of obs.	5 %	10 %	25 %	**50 %**	75 %	90 %	95 %
2004	252	12.70	13.09	14.30	**15.33**	16.55	18.13	18.91
2005	252	10.79	11.10	11.68	**12.52**	13.64	14.84	15.59
2006	251	10.55	10.79	11.36	**12.00**	13.62	16.21	17.73
2007	251	10.37	11.10	13.13	**16.43**	21.66	25.25	26.49
2008	253	18.61	19.66	21.58	**25.10**	40.00	60.86	67.70
2009	252	21.14	22.11	24.28	**28.57**	39.31	45.43	47.41
2010	252	16.47	17.29	18.34	**21.72**	25.20	29.63	29.63
2011	252	15.82	16.06	17.40	**20.72**	31.57	36.20	39.01
All	2015	10.85	11.46	13.53	**17.79**	24.86	34.21	43.81

Over the entire sample, the median of the VIX falls slightly short of 18 volatility (percentage) points, and 50% of the time the indices closed in a range of roughly 11 points (between 13.52 and 24.86), whereas excluding the extremely high and extremely low values, 90% of the observations of the VIX are found in a range of roughly 33 points. The table reveals great variation in the year-to-year normal ranges. The period of lower volatility in 2004–2007 is characterized by tighter ranges, although in 2007 the wider range speaks of the beginning of the financial crisis and the increase in volatility it entailed. In 2004–2006, the VIX closed below or slightly above the full sample median at least 95% of the time. Even during this calmer period, however, VIX levels below 10 points are unlikely. At the other end of the spectrum, the financial crisis and the market turbulence that ensued are reflected in the substantially wider ranges of 2008–2009. In 2008 in particular, even the lowest percentile is nearly twice that of the full sample. The largest year-on-year change in the percentiles occurs in 2007–2008, which would suggest a regime change in the level of volatility. Overall, the normal ranges give evidence of amplified volatility movements in times of deteriorating market conditions and of changes in the average levels of market volatility expressed by the VIX.

4.2 Econometric methodology

We use the Bai and Perron (1998; 2003a;b; 2004) (henceforth: B&P) structural break method to test for multiple breaks in the VIX series at *a priori* unknown times. This method has been applied to identifying regime changes in market volatility by Guo and Wohar (2006) in their study of the old and new VIX indices. The remainder of this section presents the econometric model and the method developed by B&P for testing for multiple shifts at unknown times in the mean of a time series.

As suggested by B&P, consider the following multiple linear regression model with m breaks (corresponding to $m+1$ regimes)

$$\begin{aligned} y_t &= x_t'\beta + z_t'\delta_1 + u_t, \qquad t = 1, \ldots, T_1, \\ y_t &= x_t'\beta + z_t'\delta_2 + u_t, \qquad t = T_1 + 1, \ldots, T_2, \\ &\vdots \\ y_t &= x_t'\beta + z_t'\delta_{m+1} + u_t, \qquad t = T_m + 1, \ldots, T. \end{aligned} \tag{4.1}$$

where, at time t, y_t is the observed dependent variable, x_t is a $p \times 1$ vector of regressors, z_t is a $q \times 1$ vector of regressors and β and $\delta_j (j = 1, \ldots, m+1)$ are the correspoding vectors of regression coefficients. u_t is the error term. The regime index, also called the m-partition, $(T_1, \ldots, T_m)$ represents the set of breakpoints for the different regimes, and $T_0 = 0$ and $T_{m+1} = T$ by convention. All breakpoints are explicitly treated as unknown, and for $i = 1, \ldots, m$, we have break factions defined as $\lambda_i = T_i/T$ with $0 < \lambda_1 < \cdots < \lambda_m < 1$. The possible breakpoints are restricted to be asymptotically distinct and bounded from the boundaries of the sample. To this effect, B&P define the following set,

$$\Lambda_\epsilon = \{(\lambda_1, \ldots, \lambda_m); |\lambda_{i+1} - \lambda_i| \geq \epsilon, \lambda_1 \geq \epsilon, \lambda_m \leq 1 - \epsilon\}, \tag{4.2}$$

for some arbitrary positive number $\epsilon = h/T$, which acts as a trimming parameter by imposing a minimum number of observations h for a regime. The B&P method for identifying structural breaks is convenient in that it allows for quite general specifications in the regression model. No restrictions are imposed on the variance of the error term u_t, which allows for accommodating autocorrelation and heteroskedasticity in the residuals. Moreover, structural breaks in variance are also permitted if they occur on the same dates as the breaks in the regression parameters.

The multiple linear regression system given above is the most general version of the model considered by B&P. It is a partial structural change model since the parameter vector β is not subject to shifts over time, but is estimated using the entire sample. A special case is when $p = 0$ and the terms $x_t'\beta$ disappear. This is a pure structural change model, since all the (remaining) coefficients are subject to change. Here, we follow Bai and Perron (2003a) and Guo and Wohar (2006) in regressing the VIX time series on a constant only and test for structural breaks in that constant. Such a regression model with m breaks ($m + 1$ regimes) is given by

$$\mathrm{VIX}_t = \delta_j + u_t, \quad t = T_{j-1} + 1, \ldots, T_j, \quad j = 1, \ldots, m + 1, \tag{4.3}$$

where VIX_t is the VIX index level on day t and δ_j is the VIX mean level in the jth regime. The objective is to estimate the breakpoints $(T_1, \ldots, T_m)$ together with the unknown regression coefficients when T observations on VIX_t are available. The method of estimation, as proposed by B&P, is based on the ordinary least squares (OLS) principle. For each m-partition $(T_1, \ldots, T_m)$, denoted $\{T_j\}$, the associated OLS estimates of δ_j are obtained by minimizing the sum of squared residuals

$$S_T(T_1, \ldots, T_m) = \sum_{i=1}^{m+1} \sum_{t=T_{i-1}+1}^{T_1} (\mathrm{VIX}_t - \delta_i)^2 \tag{4.4}$$

under the constraint that $\delta_i \neq \delta_{i+1} (1 \leq i \leq m)$. Denote the resulting estimates as $\hat{\delta}(\{T_j\})$. Substituting these into the objective function (4.4) gives an expression for the estimated breakpoints $(\hat{T}_1, \ldots, \hat{T}_m)$ such that

$$(\hat{T}_1, \ldots, \hat{T}_m) = \underset{(T_1, \ldots, T_m)}{\arg\min}\, S_T(T_1, \ldots, T_m), \tag{4.5}$$

where the minimization is taken over all partitions $(T_1, \ldots, T_m)$ such that $T_i - T_{i-1} \geq \mathrm{int}(T\epsilon) \geq q$. Therefore, the breakpoint estimators (estimated dates on which the breaks occur) are obtained as the global minimum of the sum of squared residuals. The final regression parameters are then the OLS estimates associated with the m-partition $\{\bar{T}_j\}$, i.e. $\hat{\hat{\delta}} = \hat{\delta}\left(\{T_j\}\right)$. An efficient algorithm for solving the minimization problem is given in Bai and Perron (2004).

In order to identify the number of structural breaks m in the regression model (4.3), B&P consider certain testing procedures. The starting point is testing the hypothesis of no structural breaks, i.e. the case where the regression coefficients remain constant over time,

$$H_0: \quad \delta_j = \delta_0, \qquad j = 1, \ldots, m+1, \tag{4.6}$$

against the alternative hypothesis of a known number m of breakpoints, where the regression coefficients shift from one stable relationship to a different one. To this end, B&P employ a $\sup F$ type test (see Andrews 1993) of structural stability $(m = 0)$ versus $m = k$ breakpoints. Let $(T_1, \ldots, T_k)$ be a partition such that $T_i = \text{int}(T\lambda_i)$ with $i = 1, \ldots, k$. For this partition, the F-statistic for testing $\delta_1 = \cdots = \delta_{k+1}$ against $\delta_i \neq \delta_{i+1}$ is given by

$$F_T^*(\lambda_1, \ldots, \lambda_k; q) = \frac{1}{T}\left(\frac{T - (k+1)q - p}{kq}\right)\hat{\delta}' R' (R\hat{V}(\hat{\delta})R')^{-1} R\hat{\delta}, \tag{4.7}$$

where R is a matrix such that

$$(R\delta)' = (\delta_1' - \delta_2', \ldots, \delta_k' - \delta_{k+1}')$$

and $\hat{V}(\hat{\delta})$ is an estimate of the variance-covariance matrix of $\hat{\delta}$ that is robust to serial correlation and heteroskedasticity. The $\sup F$ type test statistic is then defined as

$$\sup F_T^*(k; q) = \sup_{(\lambda_1, \ldots, \lambda_k) \in \Lambda_\epsilon} F_T^*(\lambda_1, \ldots, \lambda_k; q). \tag{4.8}$$

An asymptotically equivalent, yet computationally simpler version is given by

$$\sup F_T(k; q) = F_T^*(\hat{\lambda}_1, \ldots, \hat{\lambda}_k; q), \tag{4.9}$$

where the break faction estimates $\hat{\lambda}_i = \hat{T}_i/T, i = 1, \ldots, k$ minimize the global sum of squared residuals in Equation (4.4). (Bai and Perron 2003a).

Notice that the $\sup F$ test presented above requires pre-specifying the number of structural breaks k in the hypothesis against which the null hypothesis of structural stability is tested. To allow for testing the null hypothesis against a more vague alternative hypothesis of an unknown number of breakpoints m given an upper bound M, B&P develop two tests, referred to as the *double maximum tests*. For some fixed weights $\{a_1, \ldots, a_m\}$, the tests are given by

$$D\max F_T^*(M, q, a_1, \ldots, a_m) = \max_{1 \leq m \leq M} a_m \sup_{(\lambda_1, \ldots, \lambda_m) \in \Lambda_\epsilon} F_T^*(\lambda_1, \ldots, \lambda_m; q) \tag{4.10}$$

Again, an asymptotically equivalent and computationally simpler version exists and is given by

$$D\max F_T(M, q, a_1, \ldots, a_m) = \max_{1 \leq m \leq M} a_m F_T(\hat{\lambda}_1, \ldots, \hat{\lambda}_m; q). \tag{4.11}$$

The version of the test, labeled $UD\max F_T(M, q)$, is where all weights are equal to unity, $a_1 = \cdots = a_m = 1$. For the second test, denoted $WD\max F_T(M, q)$ the set of weights is such that the marginal p-values are equal across numbers of breakpoints, i.e. different values of m. The weights are defined as $a_1 = 1$ and $a_m = c(q, \alpha, 2)/c(q, \alpha, m)$ for $m > 1$, where α is the significance level of the test and $c(q, \alpha, m)$ is the asymptotical critical value of the sup F test in (4.8).[1] Thus, we have

$$UD\max F_T(M, q) = \max_{1 \leq m \leq M} F_T(\hat{\lambda}_1, \ldots, \hat{\lambda}_m; q) \tag{4.12}$$

and

$$WD\max F_T(M, q) = \max_{1 \leq m \leq M} \frac{c(q, \alpha, 1)}{c(q, \alpha, m)} F_T(\hat{\lambda}_1, \ldots, \hat{\lambda}_m; q). \tag{4.13}$$

It should be noted that unlike the $UD\max F_T(M, q)$ test, the value of the $WD\max F_T(M, q)$ depends on the level of significance chosen by the researcher. An estimated test statistic above the specified critical values in of these tests suggests rejection of the null hypothesis of no structural breaks in favor of at least one, but no more than M breaks.

In order to determine the exact number of breakpoints in the data, B&P discuss the use of the well-known Bayesian Information Criterion (BIC) and, as an alternative, a modified Schwarz' criterion (Liu, Wu, and Zidek 1997) in estimating the number of breakpoints. Based on a simulation study by Perron (1997), they conclude on a number of shortcomings related to the information criteria's performance in the presence of serial correlation and a lagged dependent variable in the error term u_t, which can lead to biased estimates

[1] Critical values for differents values of M and the trimming value ϵ are provided in Bai and Perron (2003b).

for the number of breaks.

As an alternative to convenational information criteria, B&P develop a test for the null hypothesis of ℓ breaks versus the alternative hypothesis of $\ell+1$ breaks in order to ascertain whether an additional break leads to a statistically significant reduction in the sum of squared residuals, i.e. whether an additional break exists or not. This test is labelled $\sup F_T(\ell+1|\ell)$. It is a sequantial test inasmuch as it applies $\ell+1$ tests of the null hypothesis of structural stability against the alternative hypothesis of a single structural change on a model with ℓ breakpoints. The test is applied to each of the $\ell+1$ regimes containing the observations $\hat{T}_{i-1}+1,\ldots,\hat{T}_i$ for $i=1,\ldots,\ell+1$ and remembering that again $\hat{T}_0=0$ and $\hat{T}_{\ell+1}=T$ by convention. The null hypothesis is rejected in favor of a model with $\ell+1$ breaks if the overall minimal value of the sum of squared residuals is sufficiently smaller than that of the model with ℓ breaks. The additional breakpoint is the one associated with this overall minimal value. Formally, we have

$$F_T(\ell+1|\ell) = \frac{1}{\hat{\sigma}^2}\left\{S_T\left(\hat{T}_1,\ldots,\hat{T}_\ell\right) - \min_{1\le i\le \ell+1}\inf_{\tau\in\Lambda_{i,\eta}} S_T\left(\hat{T}_1,\ldots,\hat{T}_{i-1},\tau,\hat{T}_i,\ldots,\hat{T}_\ell\right)\right\}, \tag{4.14}$$

where

$$\Lambda_{i,\eta} = \left\{\tau;\hat{T}_{i-1}+\left(\hat{T}-\hat{T}_{i-1}\right)\eta \le \tau \le \hat{T}_i - \left(\hat{T}-\hat{T}_{i-1}\right)\eta\right\}$$

and $\hat{\sigma}^2$ is a consistent estimate of the model's sample variance σ^2 under the null hypothesis. The sequential application of this testing procedure begins obviously with the $\sup F_T(1|0)$ statistic. However, Bai and Perron (2004) argue that this initial statistic can have low power in the presence of multiple breaks. Therefore, they recommend first examining the double maximum test statistics to determine whether any structural breaks are present. In case the statistics are statistically significant, the researcher should then examine the higher order $\sup F_T(\ell+1|\ell)$ statistics to decide on the exact number of breaks, by choosing the highest value of ℓ for which the statistic is rejected. When allowing for heteroskedasticity and serial correlation in the data series under scrutiny, the authors recommend using a trimming parameter of at least $\epsilon=0.15$, which corresponds to an upper bound on the number of structural breaks $M=5$.

In our application below, we follow these guidelines and allow for heteroskedasticity and serial correlation in the regression residuals (using the notation of Bai and Perron (2004), we set $cor\_u=1$, $het\_u=1$). Since our sample period consists of $T=2015$ observations (trading days), the trimming parameter $\epsilon=0.15$ imposes a maximum length of

$h = T\epsilon = 302$ trading days on each regime. The estimation procedure is carried out using a MATLAB version of B&P's original GAUSS code developed by Yohei Yamamoto.[2]

4.3 Results

The results for B&P statistics for tests of structural change in the mean value of the VIX series are given in Table 4.2. The table also displays the break locations obtained from the global optimization of models with $m = 1, \ldots, 5$ breaks. Both double maximum test statistics ($UD\max$ and $WD\max$) are clearly statistically significant at all levels of significance, which strongly suggests the presence of one or more structural breaks in the mean level of the VIX. The initial $\sup F(\ell + 1|\ell)$ test of zero structural breaks against a single break rejects the null hypothesis, supporting the evidence presented by the double maximum tests. Such is the case also with the $\sup F(3|2)$ test statistic, which also clearly points to rejection of the null hypothesis of two breaks against the alternative hypothesis of three breaks. However, for the $\sup F(2|1)$ test, the estimated statistic is not statistically significant, which gives further evidence of only one structural break in the data. This leads to the conclusion that there is a single structural break which divides the VIX time series into two regimes.

Table 4.2 also gives the breakpoint locations estimated by global optimization of model (4.5) with $m = 1, \ldots, 5$. For the model with a single structural break, the breakpoint is placed on July 24, 2007. This coincides approximately with the time when the consequences of the U.S. subprime mortgage crisis began to unfold and spread into the rest of the world. Note that the bankruptcy of Lehman Brothers, which triggered levels of volatility second only to the stock market crash of October 1987 (the "Black Monday"), is visible in the estimated break dates of the models with 2, 3 and 4 breaks, all of which place one break to early September, 2008.

Based on the break date estimate, the VIX time series can be separated into two different volatility regimes. Regime 1 extends from 1 January 2004 to 24 July 2007 and Regime 2 spans the latter part from 25 July 2007 to 31 December 2011. Figure 4.2 gives a visual depiction of the regime change in the VIX time series. Indeed, the latter part of the sample period exhibits notably higher levels of volatility, and the volatility of volatility itself seems higher as well. The mean level of the VIX is 13.67, and any deviations thereof

[2]Both versions are available at `http://people.bu.edu/perron`.

Table 4.2: Bai and Perron structural break test results for the VIX series, sample period 02/01/2004–31/12/2011.

	Test statistic	Break dates				
		$\hat{T}_1$	$\hat{T}_2$	$\hat{T}_3$	$\hat{T}_4$	$\hat{T}_5$
$UD\max$[a]	39.89***					
$WD\max(5\%)$[b]	57.42***					
$F(1\|0)$[c]	19.27***	24/07/2007				
$F(2\|1)$[f]	8.10	02/09/2008	11/11/2009			
$F(3\|2)$[e]	87.74***	26/06/2007	05/09/2008	16/11/2009		
$F(4\|3)$[f]	9.39	16/05/2005	26/06/2007	05/09/2008	16/11/2009	
$F(5\|4)$[g]	-	16/05/2005	09/03/2007	28/05/2008	07/08/2009	19/10/2010

[a] Test of the null hypothesis of 0 breaks against the alternative hypothesis of an unknown number of breaks given an upper bound of $m = 5$. The 10%, 5% and 1% critical values equal 7.46, 8.88 and 12.37, respectively.
[b] Test of the null hypothesis of 0 breaks against the alternative hypothesis of an unknown number of breaks given an upper bound of $m = 5$. The 5% critical value equals 9.91.
[c] Test of the null hypothesis of $\ell = 0$ breaks against the alternative hypothesis of $\ell + 1 = 1$ break. The 10%, 5% and 1% critical values equal 7.04, 8.58 and 12.29, respectively.
[d] The 10%, 5% and 1% critical values equal 8.51, 10.13 and 13.89, respectively.
[e] The 10%, 5% and 1% critical values equal 9.41, 11.14 and 14.80, respectively.
[f] The 10%, 5% and 1% critical values equal 10.04, 11.83 and 15.28, respectively.
[g] The 10%, 5% and 1% critical values equal 10.58, 12.25 and 15.76, respectively.
***, ** and * represent statistical significance at the 10%, 5% and 1% levels of confidence.
- indicates that given the location of the breaks there was no more place to insert an additional break that would satisfy the minimal length requirement (here $h = T\epsilon = 302$).

are relatively small. The mean level of volatility in the second regime is 27.24, which is more than twice that of the first regime, largely as a result of the several volatility peaks from late 2008 onwards. Standard deviation is found to be 2.28 in Regime 1 and 11.00 in Regime 2; again, the increase is considerable and speaks of higher dispersion in the volatility itself. Overall, the estimated breakpoint marks a transition from a calm period into one with higher uncertainty, and divides the VIX time series into a period of low volatility (Regime 1) and a period of high volatility (Regime 2).

In light of the results obtained above, we move on to the study of the relationship between volatility and stock market returns in and across the two volatility regimes.

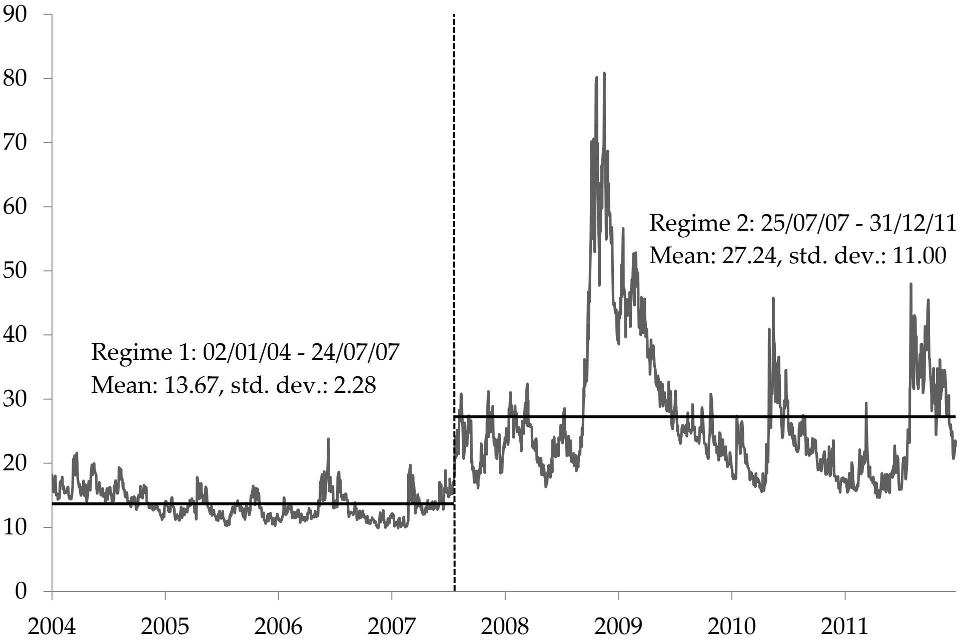

Figure 4.2: Volatility regimes and their mean levels in the VIX daily time series from January 2004 to December 2011.

Chapter 5

Relationships between implied volatility and stock market returns

In what follows we examine the interactions between daily returns on the underlying stock index and daily changes in their implied volatility, as measured by daily differences in the VIX. In particular, we try to model their contemporaneous relationship and look for any evidence of asymmetries therein. Furthermore, we examine the direction of causality in the relationship, and attempt to quantify the predictive abilities of the VIX and the S&P 500 on each other.

The S&P 500 is a stock market index that is calculated from the market capitalizations of 500 leading, publicly traded companies in the United States stock market. It is one of the most commonly followed equity indices even internationally, and is widely considered as the best indicator of market sentiment and a bellwether for the United States economy. As an index calculated from the options on the S&P 500, the VIX is often considered the foremost measure of the market's expectation of stock market volatility. The VIX has earned the moniker "fear index" or "fear gauge", since it tends to react more strongly to bearish events, i.e. events that have a negative impact on stock markets. Such events are associated with deteriorating investor sentiment and increasing risk, of which the VIX can be considered a metric.

The daily closing values for the VIX and its underlying stock index, the S&P 500 (henceforth referred to as the SPX, according to its ticker symbol), are obtained from Yahoo! Finance for a period from January 2, 2004 to December 30, 2011, spanning 2015 trading days. We calculate the SPX returns in the standard way as the log (continuously compounded) return rather than arithmetic return[1] given by

$$R_t = 100 \ln \frac{\mathrm{SPX}_t}{\mathrm{SPX}_{t-1}}, \tag{5.1}$$

where SPX_t denotes the SPX index level on day t. The log ratio is multiplied by 100 to give the return an interpretation as percentages and thereby to facilitate comparison with the VIX index values, which are given in volatility units (percentage points).

5.1 Summary statistics

Table 5.1 presents the summary statistics of daily VIX and SPX levels as well as VIX changes and SPX returns for the entire sample period 2004-2011 and the two previously identified volatility regimes. Also included are the test statistics for testing whether the data are normally distributed as well as for testing for the presence of unit roots in the time series. The VIX ranges from a minimum of 9.89 % to a maximum of 80.86 %, its highest value on November 2, 2008. Some time thereafter, the SPX tumbled to its in-sample minimum of less than 700 points, a far cry from the highest level reached in the beginning of Regime 2. The VIX levels and especially VIX changes as well as SPX returns exhibit sizeable excess kurtosis, and the VIX levels and changes are both positively skewed. The VIX and SPX levels both show highly persistent first-order autocorrelation (given by the coefficient ρ_1), which is not removed by transforming the series into first differences and first log differences, respectively.

JB stands for the Jarque and Bera (1987) test statistic and is given by

$$JB = \frac{n}{6}\left(\mathrm{Skewness}^2 + \frac{1}{4}(\mathrm{Kurtosis} - 3)^2\right), \tag{5.2}$$

where n is the sample size. It is a goodness-of-fit test of whether the sample skewness and

[1] Arithmetic return, the conventional definition of the rate of change, is given by $R_{t,arithm} = (S_t - S_{t-1})/S_{t-1}$. Therefore, the return from holding an asset for one period from time $t-1$ to time t is $1 + R_{t,arithm} = S_t/S_{t-1}$. This return becomes logarithmic under continuous compounding:

$$\lim_{m\to\infty}(1 + R_{t,arithm})^{\frac{1}{m}} = \ln(1 + R_t) = \ln\left(\frac{S_t}{S_{t-1}}\right) = \ln S_t - \ln S_{t-1}.$$

Provided that the daily changes remain small, logarithmic return is approximately equal to arithmetic return:

$$\ln(1 + R_{t,arithm}) \approx R_{t,arithm}, \quad R_{t,arithm} \ll 1,$$

and thus it can be used as a proxy for rate of change over one period.

kurtosis match those from a normal distribution; a statistically significant value leads to rejection of this null hypothesis in favor of an alternative probability distribution. Clearly, the null hypothesis of normally distributed data is strongly rejected for the VIX and SPX levels as well as VIX changes and SPX returns.

The non-normality of observed asset price returns is contradictory to the theoretical asset price processes discussed in Chapter 2. For example, if an asset price process S_t is assumed to follow a geometric Brownian motion, Equations (2.8) and (2.9) show its logarithmic returns (defined analogously to Equation (5.1)) are normally distributed. However, empirical evidence against the assumption of normally distributed asset returns has been mounting ever since the studies of Mandelbrot (1963), Fama (1965) and Clark (1973). Instead, stock returns tend to have leptokurtic distributions, i.e. distributions that are more peaked around the mean and have fatter tails[2] than a normal distribution. This implies that extremely unlikely observations, i.e. extremely high returns of either signs, are more likely than a normal distribution suggests.[3] More recent findings argue in favor of a (scaled) Student's t-distribution; see e.g. Peiró (1994), Egan (2007) and Platen and Rendek (2008). Note that the concept of model-free implied volatility is valid with non-normal asset returns, as the theory makes no assumptions on the specific process governing the underlying asset returns (see e.g. Britten-Jones and Neuberger 2000).

Looking again at Table 5.1, ADF stands for the Augmented Dickey-Fuller test[4], which is a test for a unit root in a time series sample. The null hypothesis is that there is a unit root in the sample, while the alternative hypothesis suggests that the series is stationary. The KPSS (Kwiatkowski, Phillips, Schmidt, and Shin 1992) is an alternative test for stationarity. It is included here in order to complement the results of the ADF test. The KPSS test tests a null hypothesis of stationarity around a deterministic trend against an alternative hypothesis of a unit root in the sample. According to the results, both tests give strong evidence of a unit root in both the VIX and the SPX level time series, however, VIX changes and SPX returns appear to be stationary.

[2]Fat tails in a probability distribution mean that extreme values have higher probability density and are thus more likely to occur. An asset whose returns are characterized by a leptokurtic distribution is therefore more risky than an asset with normally distributed returns would be.

[3]Extremely unlikely events that are highly unpredictable and fall beyond "regular" expectations are sometimes referred to as Black Swans (see Taleb 2001).

[4]An augmented version of the original test developed by Dickey and Fuller (1979).

Table 5.1: Descriptive statistics of the VIX and the S&P 500 levels and returns for the full sample period and Regimes 1 and 2.

	VIX		S&P 500	
	Levels	Changes	Levels	Returns (%)
Full sample (02/01/2004–31/12/2011), N=2015				
Mean	21.21	0.00	1212.48	0.01
Std. dev.	1.99	1.99	168.84	1.39
Min.	9.89	-17.36	676.53	-9.47
Max.	80.86	16.54	1553.08	10.96
Kurtosis	8.18	20.25	3.16	12.85
Skewness	2.03	0.55	-0.37	-0.30
$\rho(1)$	0.98***	-0.16***	0.99***	-0.12***
JB	3634.34***	25082.55***	50.49***	8167.04***
ADF	-3.04	-13.16***	-1.72	-12.17***
KPSS	5.68***	0.03	2.14***	0.09
Regime 1 (02/01/2004–24/07/2007), N=896				
Mean	13.67	-0.00	1255.97	0.04
Std. dev.	2.28	0.87	121.52	0.67
Min.	9.89	-5.66	1063.23	-3.53
Max.	23.81	7.16	1553.08	2.13
Kurtosis	3.35	12.05	2.56	4.06
Skewness	0.75	0.73	0.70	-0.25
$\rho(1)$	0.92***	-0.10***	0.99***	-0.05
JB	89.59***	3132.87***	79.88***	51.01***
ADF	-2.88	-11.24***	-2.57	-10.49***
KPSS	2.78***	0.09	11.38***	0.06
Regime 2 (25/07/2007–31/12/2011), N=1119				
Mean	27.24	0.00	1177.65	-0.02
Std. dev.	11.00	9.84	191.84	1.76
Min.	14.62	-17.36	676.53	-9.47
Max.	80.86	16.54	1565.15	10.96
Kurtosis	7.00	13.34	2.45	8.82
Skewness	1.89	0.45	-0.29	-0.22
$\rho(1)$	0.97***	-0.16***	0.99***	-0.13***
JB	1415.17***	5030.09***	1177.96***	1591.73***
ADF	-2.54	-10.31***	-1.55	-10.36***
KPSS	1.54***	0.04	2.60***	0.24

This table reports descriptive statistics for VIX daily levels and changes (first differences) as well as for S&P 500 daily levels and returns. $\rho(k)$ denotes the k-th order autocorrelation. *, ** and *** denote statistical significance at the 10%, 5% and 1% levels, respectively. Statistical significance of the Jarque-Bera (JB) test leads to rejection of the null hypothesis that the observed skewness and kurtosis match those of a normal distribution. Statistical significance of the ADF test statistic suggests rejection of a null hypothesis of a unit root in the time series, whereas statistical significance of the KPSS test statistic suggests rejection of a null hypothesis that the time series is stationary around a deterministic trend.

The summary statistics are particularly revealing when examined in the two volatility regimes. As was already demonstrated in Figure 4.2, the mean and standard deviation of VIX levels are significantly higher in the high-volatility Regime 2 than in Regime 1, a period of lower volatility; the average VIX level in Regime 2 is approximately double the mean in Regime 1, whereas the standard deviation increases to nearly fivefold in Regime 2. The ranges of VIX values were 66.24 percentage points in Regime 2 and a mere 13.92 percentage points in Regime 1, and the most extreme movements in the VIX in either direction occurred during Regime 2. This is also the case with SPX returns, where the extreme observations amount to approximately one tenth of the index value. The mean of SPX returns shifted from marginally positive in Regime 1 to marginally negative in Regime 2, while the mean of VIX changes eked to positive territory in Regime 2. Note that in Regime 1, the first-order autocorrelation in SPX returns is not statistically significant.

Daily levels for the VIX and the underlying SPX are shown in Figure 5.1. The graph suggests that the changes in implied volatility and the stock index returns are negatively correlated at least to some degree: increases in the stock index level would appear to correspond to decreases in implied volatility, and vice versa. This inverse relation is perhaps somewhat more clearly visible in the latter volatility regime, where the overall movement is higher in magnitude. However, the relationship would seem to persist throughout the sample period.

Note that the VIX is characterized by higher fluctuations and notable upward spikes compared to SPX returns. Most of these spikes are coincident with spikes in the opposite direction for the underlying stock indices at times when potentially risky events have taken place – volatility increases more when stocks decline than it decreases when stocks go up. The SPX trended up over most of Regime 1, whereas the corresponding downtrend in the VIX is notably milder or nearly nonexistent. Infact, it is as if the VIX has a strong support level at around 10 volatility points. In Regime 2, particularly noteworthy is the beginning of the subprime crisis and subsequent global financial crisis in late 2008, when implied volatility skyrocketed to its all-time highest level, while the stock indices plummeted to their lowest levels in more than a decade. More recently, the worrying turns of events resulting in the ever-ongoing European sovereign debt crisis are represented by spikes in volatility.

The upward spikes relating to worriesome events and market turmoil are the reason why the VIX is sometimes dubbed an ”investor fear gauge”. As Whaley (2009) points out, a more elaborate explanation arises from the fact that the demand for options increases when investors become concerned about a drop in the stock market and seek to hedge risk. The increased demand of options then raises their prices, which in turn are transferred to the volatility indices resulting in higher levels of volatility. As investor fear subsides, option prices tend to decline, causing in turn the VIX to decline – however, the decreases in volatility relating to favorable market conditions are visibly milder, as can be seen from Figure 5.1 below.

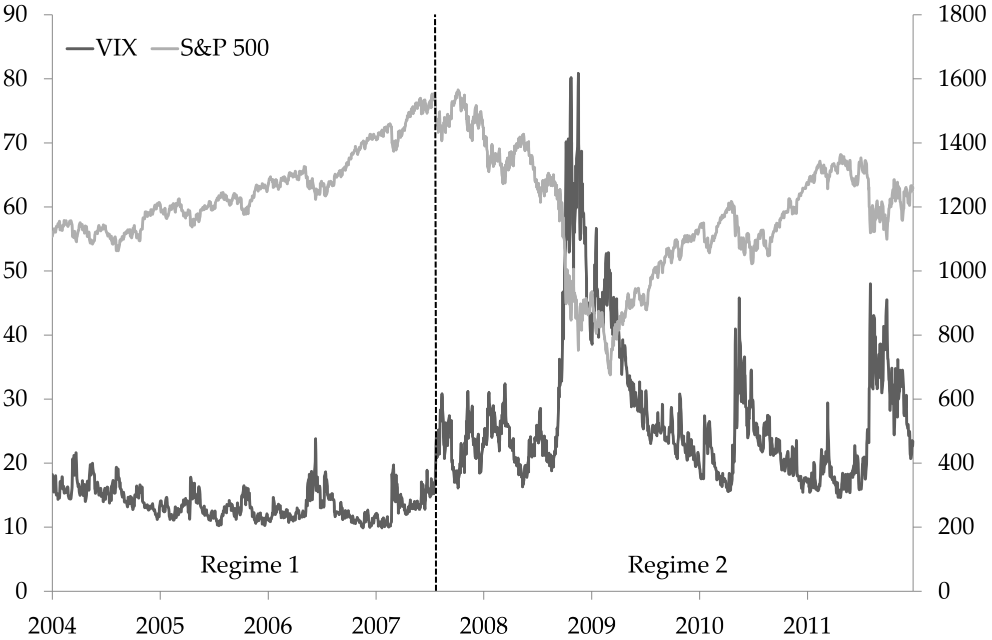

Figure 5.1: Daily closing values of the VIX volatility index (left scale) and the S&P 500 stock index (right scale) from January 2004 to December 2011.

5.2 Intertemporal cross-correlations

In order to more accurately analyze the inverse relationship between VIX changes and SPX returns, we estimate the intertemporal cross-correlations between lags and leads on asset returns and changes in implied volatility by using the standard definition of sample cross-correlation between two time series X_t and Y_t both consisting of n observations x_i and y_i:

$$\rho_{XY}(k) = \frac{\gamma_{XY}(k)}{s_X s_Y}, \tag{5.3}$$

where k is the number of lags, $s_X = \sqrt{\gamma_{XX}(0)}$ and $s_Y = \sqrt{\gamma_{YY}(0)}$ are the sample standard deviations. The sample cross-covariance $\gamma_{XY}(k)$ is given by

$$\gamma_{XY}(k) = \begin{cases} \frac{1}{n}\sum_{i=1}^{n-k}(x_i - \bar{x})(y_{i+k} - \bar{y}), & k = 0, 1, 2, \ldots \\ \frac{1}{n}\sum_{i=1}^{n+k}(x_i - \bar{x})(y_{i-k} - \bar{y}), & k = 0, -1, -2, \ldots \end{cases} \tag{5.4}$$

where $\bar{x} = \sum_{i=1}^{n} x_i$ and $\bar{y} = \sum_{i=1}^{n} y_i$ are the sample means. Figure 5.2 plots the cross-correlation estimates between SPX returns at different positive and negative lags k against changes in the VIX in regimes 1 and 2. The two dashed lines denoting the 95% confidence interval in both graphs. Indeed, the contemporaneous correlations on any given day t are strongly negative at -0.80 and -0.85 in regimes 1 and 2, respectively. The strongly negative contemporaneous correlations are evidence of the leverage effect, as proposed by Black (1976) and Christie (1982), who argued that given a fixed debt level, a decline in the equity level increases the leverage of a company (market) and hence the risk for its stock (index). The increased risk translates to an increase in stock (index) volatility.[5]

At other lags ($k \neq 0$) the correlations are notably smaller. In Regime 1, the first two negatively lagged returns ($k = -1, -2$) show marginally significant positive correlations with changes in the VIX, indicating that SPX returns have some predictive power over future movements in the VIX. In Regime 2 also the correlations of the first two positively lagged returns become statistically significantly positive, which suggests that movements in the VIX would have some predictive power over future SPX returns *and* vice versa. Our findings on the correlations for Regime 1 are consistent with Carr and Wu (2006) – however, in Regime 2 the intertemporal correlations seem to run both ways: The SPX returns with both positive and negative lags and leads show (marginally) statistically significant correlations with VIX changes.

The study of cross correlations undertaken above suggests changes in the relationship between stock market returns and volatility relative to the observed mean shift in volatility. The latter volatility regime represents a period of abnormal market conditions

[5] The degree of correlation is generally found to be much larger for stock indices than individual stocks (see. e.g. Figlewski and Wang 2000).

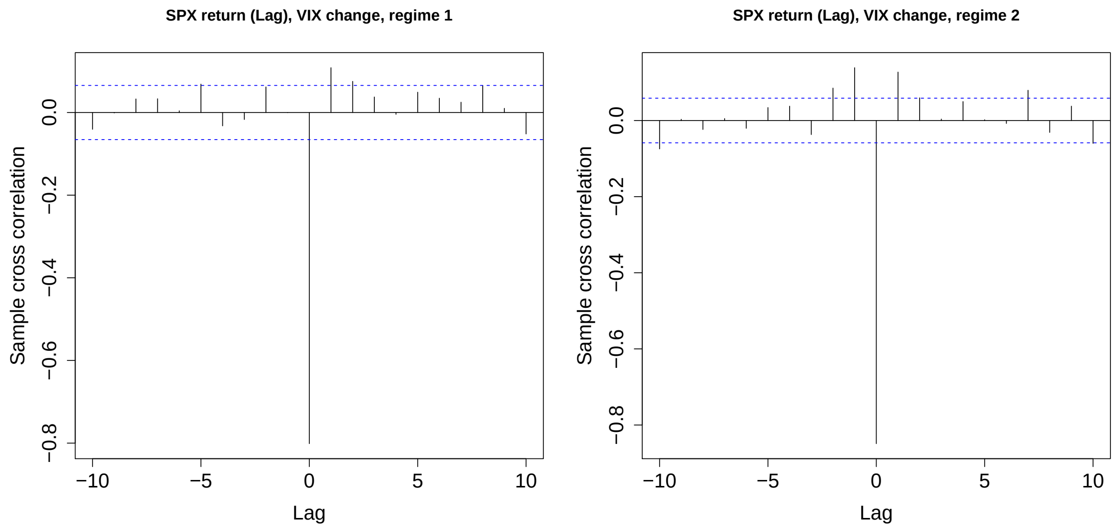

Figure 5.2: Sample cross correlations between lagged SPX returns and changes in the VIX in Regimes 1 and 2.

with unforeseen levels of and changes in volatility and could therefore be a source of obscurity. What's more, the intertemporal cross-correlations give mixed evidence on the possible predictive power of changes in volatility over future asset returns and vice versa. In what follows, we examine the possible causal links in their relationship more closely.

5.3 Granger causality tests

One way to test whether implied volatility foreshadows returns (or vice versa) is to perform Granger (1969) causality tests. A time series Y is said to *Granger cause* X if it can be shown that lagged values of Y provide statistically significant information about future values of X. The notion of Granger causality thus refers to a narrower interpretation than what is commonly understood by causality: If a more accurate forecast of the future values of X can be obtained using information from the lagged values of Y along with lagged values of X than using only the lagged values of X, then Y contains useful information in forecasting future values of X or, in other words, Y Granger causes X. Conversely, Y fails to Granger cause X if for all $s > 0$ the mean squared error[6] of a fore-

[6]Mean squared error (MSE) is a common measure of the difference between values implied by an estimator and the true values being estimated. It is defined as $\frac{1}{n}\sum_{i=1}^{n}(\hat{Y}_i - Y_i)^2$, where $\hat{Y}_i$ are predictions and Y_i represent the true values ($i = 1, \ldots, n$).

cast x_{t+s} based on $(x_t, x_{t-1}, \dots)$ is equal to the mean squared error of a forecast x_{t+s} that uses both $(x_t, x_{t-1}, \dots)$ and $(y_t, y_{t-1}, \dots)$ (Hamilton 1994).

The standard Granger test for causality, then, consists of running bivariate regressions of the form

$$x_t = \sum_{i=1}^{k} \alpha_i x_{t-i} + \sum_{i=1}^{k} \beta_i y_{t-i} + u_t \tag{5.5}$$

where α_i and β_i are regression coefficients, u_t is the random error term and $t = 1, \dots, T$. The lag length k is chosen more or less arbitrarily, although it can be determined using model selection criteria. Under the null hypothesis $H_0 : \beta_1 = \dots = \beta_k = 0$, Y fails to Granger cause X. The hypothesis is tested by conducting an F-test, where the unrestricted regression is defined by equation (5.5), whereas the restricted regression will only include lags of the dependent variable:

$$x_t = \sum_{i=1}^{k} \gamma_i x_{t-i} + v_t, \tag{5.6}$$

where γ_i ($i = 1, \dots, T$) are regression coefficients and v_t is the error term. Under the null hypothesis, the F-statistic

$$F = \frac{\sum_{t=1}^{T} \hat{v}_t^2 - \sum_{t=1}^{T} \hat{u}_t^2}{\sum_{t=1}^{T} \hat{u}_t^2} \frac{T - 2k - 1}{k}, \tag{5.7}$$

where $\hat{u}_t$ and $\hat{u}_t$ are the residuals from the regressions, has an asymptotic $F(k, T - 2k - 1)$ distribution as $T \to \infty$. Values of F exceeding the critical value of the F-distribution for a specified confidence level lead to rejection of the null hypothesis.

He and Maekawa (2001) assert that testing for Granger causality between two time series in the conventional way often leads to spurious causality if one or both of the time series is non-stationary. This problem can be addressed using an "augmented Granger test" developed by Toda and Yamamoto (1995) that allows for testing the causality of (co-)integrated time series. However, stationarity can often be achieved by time series transformation and since we are working with logarithmized and or differenced time series, the results from Table 5.1 confirm that the stationarity condition holds. Therefore, we test for the existence of Granger causality in either direction by running the following unrestricted regressions

$$\Delta VIX_t = \sum_{i=1}^{3} \alpha_i \Delta VIX_{t-i} + \sum_{i=1}^{3} \beta_i R_{t-i} + u_t \tag{5.8}$$

$$R_t = \sum_{i=1}^{3} \alpha_i R_{t-i} + \sum_{i=1}^{3} \beta_i \Delta VIX_{t-i} + u_t \tag{5.9}$$

where ΔVIX_t is the daily change in the VIX on day t and r_t is the SPX daily return on day t. The lag length $k = 3$ is determined using the Bayesian information criterion (BIC; see Schwarz 1978). Before running the F-tests against the restricted counterparts of the above equations, we test for heteroskedasticity and serial correlation in the residuals of the unrestricted regressions. The Breusch and Pagan (1979) test statistics for heteroskedasticity in the squared residuals of both regressions (5.8) and (5.9) (not printed here) are highly statistically significant for the entire sample period from January 2004 to December 2011 as well as for the two volatility regimes defined above, which leads to rejection of the null hypothesis of no heteroskedasticity. Similarly, the Breusch (1978)/Godfrey (1978) test statistics suggest rejection of the null hypothesis of no serial correlation in the residuals.

Since the residual diagnostics suggest the presence of both heteroskedasticity and serial correlation, we employ Newey and West (1987) heteroskedasticity and autocorrelation consistent (HAC) standard errors[7] in the Granger (1969) causality tests. Results are given below in Table 5.2. They do not give evidence of Granger causality in either direction – SPX returns don't seem to Granger cause changes in the VIX, nor is the reverse true. There is, however, some variation in the F-test statistics between the volatility regimes. The absence of Granger causality could imply that the positive correlations between positively and negatively lagged values of SPX returns and VIX changes are caused by a common third factor that drives the two variables.

Our findings contradict the results obtained by Malz (2000), who studied several measures of volatility in different markets and found statistical evidence indicating that implied volatility contains information regarding future large-magnitude returns on the underlying asset prices. In his study of a volatility index on Greek equity markets, Skiadopoulos (2004) found in turn that underlying asset returns Granger cause changes in

[7]Here, as well as in all subsequent instances where HAC standard errors are employed, the number of lags is chosen automatically using Schwarz's information criterion.

Table 5.2: Results of Granger causality tests for SPX returns and VIX changes.

Null hypothesis	F-statistic	p-value
Full sample (02/01/2004–31/12/2011), N=2015		
R does not Granger cause ΔVIX	1.61	0.184
ΔVIX does not Granger cause R	1.08	0.358
Regime 1 (02/01/2004–24/07/2007), N=896		
R does not Granger cause ΔVIX	1.52	0.201
ΔVIX does not Granger cause R	1.88	0.131
Regime 2 (25/07/2004–31/12/2011), N=1119		
R does not Granger cause ΔVIX	1.58	0.191
ΔVIX does not Granger cause R	2.91	0.406

This table reports the results for the Granger (1969) causality test for the full sample period as well as for regimes 1 and 2. *, ** and *** denote statistical significance at the 10%, 5% and 1% levels, respectively. Statistical significance of the F-statistic suggests rejection of a null hypothesis of no Granger causality.

volatility, not vice versa. Neither author mentions accounting for the assumed presence of heteroskedasticity and autocorrelation in their data. This appears to be a crucial observation – without accounting for heteroskedasticity and serial correlation by applying the Newey and West (1987) heteroskedasticity and autocorrelation consistent (HAC) standard errors, our Granger tests give results similar to those obtained by Malz (2000) (VIX changes Granger cause SPX returns). However, our final conclusion from this study remains that no evidence of statistically significant Granger causality is found.

5.4 Asymmetries in volatility

In the absence of statistically significant Granger causalities we move on to study the contemporaneous relationships between volatility and stock index returns more closely. Visual inspection of the data as well as the correlation study in Section 5.2 suggest an inverse relationship between the two, with perhaps some degree of asymmetry related to positive vis-à-vis negative changes in volatility.

The negative relationship between stock market returns and volatility has been well established by previous studies. In earlier literature, the relationship between stock index returns and corresponding changes in volatility has been found to be negative among others by Schwert (1989; 1990) and Fleming, Ostdiek, and Whaley (1995); the latter find a large negative contemporaneous correlation between VXO changes and OEX returns, suggesting an inverse relation between implied volatility and asset prices. They also report evidence of asymmetry in the relationship. Whaley (2000; 2009), Simon (2003) and Giot (2005) present similar findings in the U.S. equity markets. In Europe, analoguous relationships have been documented by e.g. González and Novales (2009) on German and Swiss stock markets and by González and Novales (2011) and Skiadopoulos (2004), who constructed artificial volatility indices for Spanish and Greek stock markets, respectively. Other studies include Dowling and Muthuswamy (2005) and Frijns, Tallau, and Tourani-Rad (2010) for evidence from Australia, Ting (2007) and Kumar (2012) for Korean and Indian markets, respectively, and finally Siriopoulos and Fassas (2009), who review twelve implied volatility indices and give results suggesting the universality of an asymmetric, negative relationship between implied volatility and asset returns.

We begin by examining the impact of "news" on implied volatility by considering the effect of a return shock on day t on same-day changes in the VIX. As the VIX represents the market's expectation of realized volatility over a 22 trading day period, ΔVIX_t can be interpreted as the expected change in market volatility over the future time period from $t + 1$ to $t + 22$. Table 5.3 reports the SPX returns partitioned into twelve strata in 0.5 percentage point intervals. For each interval, the number of observations, the mean SPX return and the corresponding mean VIX change is reported. What is of note is the considerably wider spread of observations in Regime 2. There is a visibly greater number of observations with high absolute returns in Regime 2 compared to Regime 1. Also, the mean returns are higher. This is an illustration of fat-tailed distributions; the likelihood of extreme or unlikely observations is higher, which is consistent with the overall higher level of volatility that distinguishes the two regimes from each other in the first place.

Clearly, the relationship between implied volatility and stock market returns is inverse; volatility increases in response to negative daily returns, and increases more the larger the negative return shock. Correspondingly, volatility decreases along with positive returns and decreases more the larger the positive return shock. The mean levels of ΔVIX for each SPX return interval suggest that the relationship is not entirely symmetric – volatility increases more in response to negative return shocks than it decreases when returns are positive. This feature is most clearly visible when comparing the intervals with highest absolute returns. In idle markets, when absolute returns remain small, volatility remains little changed as well. Ederington and Guan (2010) point out that this is contrary to what conventional time-series models for volatility would predict: a decline in volatility associated with stagnant markets. What's more, the findings discussed above seem to partly contradict the notion of volatility clustering, i.e. the assertion that volatile markets are followed by volatile markets and stable markets are followed by stable markets. Persistence in volatility would require that changes in volatility increase along with increases in *absolute* asset returns, regardless of whether the returns are positive or negative. According to our findings, implied volatility increases only when asset returns are negative – with positive returns, volatility tends to decrease, stabilizing the market.

Figure 5.3 graphs the mean SPX returns against corresponding VIX changes, which allows for a clearer picture of the relationship described above. Note that the deviation in the tails of the curve of Regime 1 is likely due to the scarce number of observations in those extreme intervals. Thus, they can be treated as outliers. Figure 5.3 also shows relatively little differences between the two volatility regimes and the full sample. The change in volatility associated with moderate positive returns (less than 1%) would appear to be somewhat less pronounced in Regime 1 than in Regime 2. However, the overall differences in the effect of return shocks on implied volatility appear negligible.

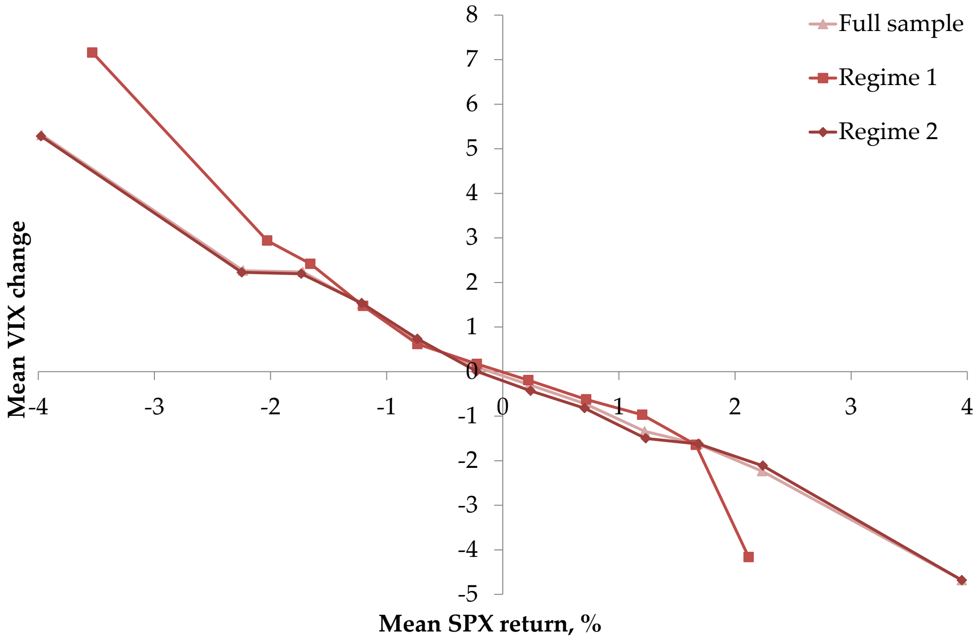

Figure 5.3: News impact curves for the VIX in the full sample and in Regimes 1 and 2.

Table 5.3: Impact of return shocks on implied volatility.

R_t (%)	# of obs.	Mean R_t (%)	Mean ΔVIX_t
Full sample (02/01/2004–31/12/2011)			
< -2.5	75	-3.97	5.31
≥ -2.5 and < -2.0	37	-2.24	2.27
≥ -2.0 and < -1.5	63	-1.72	2.23
≥ -1.5 and < -1.0	128	-1.21	1.51
≥ -1.0 and < -0.5	206	-0.74	0.68
≥ -0.5 and < 0	397	-0.22	0.10
≥ 0 and < 0.5	538	0.23	-0.30
≥ 0.5 and < 1.0	283	0.71	-0.72
≥ 1.0 and < 1.5	132	1.22	-1.34
≥ 1.5 and < 2.0	63	1.68	-1.63
≥ 2.0 and < 2.5	34	2.32	-2.23
≥ 2.5	58	3.95	-4.68
Regime 1 (02/01/2004–24/07/2007)			
< -2.5	1	-3.53	7.16
≥ -2.5 and < -2.0	2	-2.03	2.94
≥ -2.0 and < -1.5	11	-1.66	2.42
≥ -1.5 and < -1.0	50	-1.20	1.48
≥ -1.0 and < -0.5	106	-0.74	0.62
≥ -0.5 and < 0	223	-0.22	0.17
≥ 0 and < 0.5	305	0.22	-0.19
≥ 0.5 and < 1.0	139	0.72	-0.62
≥ 1.0 and < 1.5	39	1.20	-0.97
≥ 1.5 and < 2.0	16	1.66	-1.64
≥ 2.0 and < 2.5	2	2.12	-4.16
≥ 2.5	0	–	–
Regime 2 (25/07/2004–31/12/2011)			
< -2.5	74	-3.98	5.28
≥ -2.5 and < -2.0	35	-2.25	2.23
≥ -2.0 and < -1.5	52	-1.73	2.20
≥ -1.5 and < -1.0	78	-1.22	1.54
≥ -1.0 and < -0.5	100	-0.73	0.74
≥ -0.5 and < 0	174	-0.23	0.01
≥ 0 and < 0.5	233	0.24	-0.43
≥ 0.5 and < 1.0	144	0.70	-0.82
≥ 1.0 and < 1.5	93	1.23	-1.50
≥ 1.5 and < 2.0	47	1.69	-1.62
≥ 2.0 and < 2.5	32	2.24	-2.11
≥ 2.5	58	3.95	-4.68

In this table the data are partitioned by the SPX returns into twelve strata by 0.5 percentage point intervals. For each partition the number of observations and the mean levels of SPX returns as well as the corresponding mean VIX changes are reported.

Next, we further elaborate our study by exploring the above issues in a regression format. Thereby we attempt to quantify the inverse relationship between SPX returns and VIX changes and, more importantly, the degree of asymmetry in favor of volatility increases in falling markets vis-à-vis volatility decreases in rising markets. To this end, we run the following linear regressions:

$$\Delta VIX_t = \alpha_0 + \alpha_1 R_t + u_t \tag{5.10}$$

$$\Delta VIX_t = \alpha_0 + \alpha_1 R_t + \alpha_2 R_t^- + \alpha_3 \Delta VIX_{t-1} + u_t, \tag{5.11}$$

where $R_t^- = R_t$ if $R_t < 0$ and zero otherwise. Model (5.10) therefore measures simply the strength of the already observed negative relationship between returns and volatility. It is to be expected that α_1 is consistently negative and highly statistically significant. Model (5.11) adds R_t^- as a regressor to measure the level of asymmetry in the relationship; it captures the effect of negative SPX returns only. Statistical significance of the regression coefficient α_2 would be interpreted as evidence of this asymmetry. Moreover, we expect α_2 to be of the same sign as α_1, which would indicate that implied volatility increases more with respect to negative returns than it decreases with respect to positive returns. Finally, lagged VIX changes are included to check for serial correlation.

The results are reported in Table 5.4. The Newey and West (1987) heteroskedasticity and autocorrelation consistent standard errors are presented in parantheses. The intercept term (coefficient α_0) is mostly negative and significant, which suggests that if the SPX remains unchanged over the day, the VIX is expected go down, albeit only slightly. This is, of course, intuitively plausible – one would expect a measure of volatility to decrease in stagnant markets, although the observed decreased is negligible, in accordance with our earlier findings. The estimated α_0 is higher in Regime 2, reflecting the higher levels of and fluctuations in volatility during that time period. As expected, R_t is negative and highly significant as well, accounting for most of the predictive power of both models. The estimated coefficients of Model (5.10) suggest that over the full sample period, one percentage point increase (decrease) in the VIX is associated with a decrease (increase) of approximately 1.21% in the underlying SPX. In Regime 1, the SPX decreases by 1.04% for every one percentage point increase in the VIX, whereas in Regime 2 the associated daily return is approximately 1.23%. This indicates a stronger relationship in a high volatility environment. Note also that the adjusted R^2 is higher for the regressions conducted in Regime 2 than in Regime 1 – that is, SPX returns explain more of the change in implied volatility.

The estimated coefficients for R_t^- in Model (5.11) give evidence of significant asymmetry in both volatility regimes. The coefficient for ΔVIX_{t-1} is negative and statistically significant throughout, which is a sign of autocorrelation in the VIX series. The estimates for α_2 are negative, indicating that negative returns associated with positive VIX changes are higher than positive returns associated with negative VIX changes: Over the full sample period, every one percentage point increase in the VIX is followed by a $-1.0750 - 0.2360 \approx -1.31\%$ return on the SPX, whereas one percentage point decrease in the VIX translates to only a 1.075% decrease in returns. Interestingly, the coefficient α_2 measuring the degree of asymmetry is considerably higher in Regime 1, when the average level of volatility was lower. The estimated α_2 suggests a 0.47 percentage point difference in the reaction of SPX returns to positive vis-à-vis negative VIX changes; one percentage point increase in the VIX is associated with a decrease of approximately 1.27% in the SPX, while an equal decrease in the VIX is coupled with only a 0.8% increase in SPX returns. In Regime 2, the difference is only about 0.25 percentage points – however, the overall magnitude of the relationship is higher, as was already noted above in examining Model (5.10). In Regime 2, one percentage point increase in the VIX yields a 1.34% negative return on the SPX, which is somewhat higher than in Regime 1. The positive return associated with one percentage point decrease in the VIX is now 1.08%, indicating a stronger reaction than in Regime 1. This is perhaps partly explained by the strong mean reversion of the VIX, which is clearly visible in e.g. Figure 5.1: While considerably higher, the spikes in volatility are still relatively short-lived and the VIX reverts to its mean level every time, even though the SPX does not fully recover from the steep fall of 2008 over the sample period.

Our findings on the degree of asymmetry being higher during a period of lower volatility, namely Regime 1, are consistent with the results documented for the S&P 500 index returns and implied volatility by Giot (2005), who studied the relationships between the VXO (i.e. the old VIX) and its underlying stock index, the S&P 100, as well as the Nasdaq 100 stock index and its own volatility index, the VXN. Giot (2005) divided his sample period into three subperiods: one with low volatility and a bull (rising) market, one with high volatility and a bull market, and one with high volatility and a bear (falling) market. His findings indicated, similarly to ours, that volatility asymmetry was strongest during the low-volatility market environment. A possible explanation for this phenomenon is that in a regime of low volatility, investors and option traders are more sensitive to any decreases in the stock markets (negative returns) and thus more prone to hedge themselves by buying options. The increased demand for options translates to higher option prices and thereby higher implied volatility. Conversely, in a high-volatility state, option prices are already high and investors are less willing to bid them higher when the stock market falls.

Table 5.4: Regression results for Models (5.10) and (5.11).

	α_0	α_1	α_2	α_3	α_4	$\bar{R}^2$
			Full sample (02/01/2004–31/12/2011), N=2015			
Model (5.10)	-0.0101	-1.2110***				0.712
	(0.02)	(0.07)				
Model (5.11)	-0.0939**	-1.0750***	-0.2360**	-0.0769	-0.1028***	0.719
	(0.04)	(0.10)	(0.10)	(0.06)	(0.04)	
			Regime 1 (02/01/2004–24/07/2007), N=896			
Model (5.10)	-0.0365***	-1.0422***				0.638
	(0.01)	(0.07)				
Model (5.11)	-0.0837**	-0.8051***	-0.4666***	-0.0017	-0.1071***	0.660
	(0.03)	(0.09)	(0.14)	(0.05)	(0.04)	
			Regime 2 (25/07/2007–31/12/2011), N=1119			
Model (5.10)	-0.0158	-1.2310***				0.721
	(0.03)	(0.07)				
Model (5.11)	-0.1695***	-1.0820***	-0.2589**	-0.0870	-0.1031**	0.727
	(0.07)	(0.10)	(0.11)	(0.07)	(0.05)	

Newey and West (1987) heteroskedasticity and autocorrelation (HAC) consistent standard errors are given in parantheses. *, ** and *** denote statistical significance at the 10%, 5% and 1% levels, respectively.

The results documented above strongly suggest that the VIX, a measure of expected market volatility, responds more aggressively to negative changes in stock market return than to positive changes in returns of similar size. This result is consistent with the popular notion that VIX is more of a gauge of investor fear than investor positive sentiment. As Whaley (2000) puts it, the VIX therefore is rather a barometer of investors' fear of the downside than a barometer of investors' excitement in a market rally.

Chapter 6

Conclusions

In this thesis we studied the theoretical foundation, construction methodology and empirical properties of the VIX volatility index, with emphasis on identifying structural breaks in the time series data as well as their impact on the interrelationships between stock market returns and corresponding changes in implied volatility. We first presented the theoretical framework for deriving a model-free measure for volatility from option prices, and showed how the theory is implemented in the VIX index. We concentrated in the revamped calculation methodology that was introduced by the Chicago Board Options Exhange in 2003. The new calculation methodology is advantageous over its predecessor in that it is independent of any option pricing model. It extracts information from options across all available strike prices, which should make it informationally more efficient than the previous specification. Furthermore, the new methodology allows for replicating volatility as a portfolio of readily available derivative contracts, which has given way for the development of tradable products on VIX-based volatility.

In the latter part of the thesis we focused on the empirical aspects of daily VIX time series. Visual inspection of the VIX time series data suggested infrequent, but notable shifts in its mean level and the study of normal ranges of the VIX gives further reason to suspect at least one regime change in the data. We therefore used the Bai and Perron (1998; 2003a;b; 2004) method to test for structural breaks in the VIX time series. Our findings gave evidence that market volatility, as proxied by the VIX, underwent a single, significant regime shift that coincides chronologically approximately with the outbreak of the global financial crisis in late 2007. The sample time period was thereby divided into two volatility regimes corresponding to a pre-2007 period and a post 2007 period (the exact break date was found to be on 24 July, 2007), where the former regime exhibits considerably lower mean levels of volatility than the latter regime.

The study of cross-correlations confirmed the strong contemporaneous negative correlation between changes in implied volatility and stock market returns. The relationship between changes in volatility with positively and negatively lagged stock returns is somewhat vaguer, but marginally significantly positive for the nearest lags and leads. This relationship varies between the two volatility regimes. However, Granger causality tests show that no changes in volatility contain no statistically significant predictive power over future asset returns or vice versa, a result that contradicts some earlier studies. We therefore conclude that changes in volatility do not constitute a leading indicator for the stock market, nor do stock index returns hold statistically significant predictive abilities over future changes in volatility.

Upon closer examination of the contemporaneous volatility-returns relationship, the degree of the overall inverse relationship is found to be higher in Regime 2, i.e. a state of higher average volatility: positive (negative) returns associated with a unit decrease (increase) in the VIX are higher in Regime 2 than in Regime 1, and the regression models' explanatory power is higher in Regime 2. What's more, the observed inverse relationship is asymmetric: stock market implied volatility, as measured by the VIX, increases more in relation to negative stock price returns (falling markets) than it decreases in relation to positive returns (rising markets). Contrary to intuition, volatility remains roughly unchanged when returns are close to zero. Interestingly, the degree of asymmetry turns out to be higher in Regime 1, the low-volatility state, where the degree of the overall inverse relationship is lower. In summary, our findings lend further support to the VIX being an investor fear gauge.

A possible extension to this thesis would be to examine the information content and forecast quality of the VIX in the context of different volatility regimes. There exists a vast literature on forecasting volatility (see e.g. Figlewski (1997) and Poon and Granger (2003) for overviews on the subject), and of all the methods developed for this purpose, implied volatility performs generally quite well, as demonstrated by e.g. Jiang and Tian (2005). Whether this performance varies with respect to the prevalent average level of market volatility in the same manner as the degree of volatility asymmetry does, is a subject worthy of investigation. That being said, the main purpose of the VIX is not to accurately forecast the future realized volatility, but rather, it is a gauge of the expected volatility or risk currently perceived by market. As such, it is a powerful indicator of investor sentiment and provides valuable information to all market participants, which explains its popularity and the widespread use of the original VIX methodology in other markets. Furthermore, with the introduction of the new VIX methodology, volatility has become a tradable asset. The VIX offers a platform for options and futures on volatility, which provide an obvious diversification benefit owing to the strong negative correlation between stock market returns and changes in volatility.

References

S. Aboura and C. Villa. International market volatility indexes: A study on VX1, VDAX and VIX. *Available at SSRN*, 1999. URL `http://ssrn.com/abstract=394420`.

T. G. Andersen and O. Bondarenko. Construction and interpretation of model-free implied volatility. Technical report, National Bureau of Economic Research, 2007.

T. G. Andersen, O. Bondarenko, and M. T. Gonzalez-Perez. Coherent model-free implied volatility: A corridor fix for high-frequency VIX. CREATES Research Papers 2011-49, School of Economics and Management, University of Aarhus, November 2011. URL `http://ideas.repec.org/p/aah/create/2011-49.html`.

D. W. K. Andrews. Tests for parameter instability and structural change with unknown change point. *Econometrica: Journal of the Econometric Society*, pages 821–856, 1993.

L. Bachelier. *Théorie de la spéculation*. Gauthier-Villars, 1900.

J. Bai and P. Perron. Estimating and testing linear models with multiple structural changes. *Econometrica*, pages 47–78, 1998.

J. Bai and P. Perron. Computation and analysis of multiple structural change models. *Journal of Applied Econometrics*, pages 1–22, 2003a.

J. Bai and P. Perron. Critical values for multiple structural change tests. *The Econometrics Journal*, 6(1):72–78, 2003b.

J. Bai and P. Perron. Multiple structural change models: A simulation study. *Econometrics Essays*, 2004.

R. W. Banz and M. H. Miller. Prices for state-contingent claims: Some estimates and applications. *Journal of Business*, pages 653–672, 1978.

F. Black. Studies in stock price volatility changes. In *Proceedings of the 1976 meetings of the business and economic statistics section*, pages 177–181. American Statistical Association, 1976.

F. Black and M. Scholes. The pricing of options and corporate liabilities. *The Journal of Political Economy*, pages 637–654, 1973.

S. Bossu, E. Strasser, and R. Guichard. Just what you need to know about variance swaps. *JPMorgan Equity Derivatives report*, 2, 2005.

R. N. Bracewell. *The Fourier transform and its applications*. McGraw-Hill Science/Engineering/Math, 2000.

D. T. Breeden and R. H. Litzenberger. Prices of state-contingent claims implicit in option prices. *Journal of Business*, pages 621–651, 1978.

M. Brenner and D. Galai. New financial instruments for hedging changes in volatility. *Financial Analysts Journal*, 45(4):61–65, 1989.

T. S. Breusch. Testing for autocorrelation in dynamic linear models. *Australian Economic Papers*, 17(31):334–355, 1978.

T. S. Breusch and A. R. Pagan. A simple test for heteroscedasticity and random coefficient variation. *Econometrica: Journal of the Econometric Society*, pages 1287–1294, 1979.

M. Britten-Jones and A. Neuberger. Option prices, implied price processes, and stochastic volatility. *Journal of Finance*, 55(2):839–866, 2000.

P. Carr and D. Madan. Towards a theory of volatility trading. In E. Jouini, J. Cvitanic, and M. Musiela, editors, *Option Pricing, Interest Rates, and Risk Management*, pages 417–427. Cambridge University Press, 1998.

P. Carr and L. Wu. A tale of two indices. *Journal of Derivatives*, 13(3):13–29, 2006.

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

CBOE. The CBOE volatility index – VIX, 2009. URL `http://www.cboe.com/micro/vix/vixwhite.pdf`.

A. A. Christie. The stochastic behavior of common stock variances: Value, leverage and interest rate effects. *Journal of Financial Economics*, 10(4):407–432, 1982.

P. K. Clark. A subordinated stochastic process model with finite variance for speculative prices. *Econometrica*, 41(1):135–155, 1973.

J. C. Cox and S. A. Ross. The valuation of options for alternative stochastic processes. *Journal of financial economics*, 3(1):145–166, 1976.

J. C. Cox and M. Rubinstein. *Options markets*. Prentice Hall, 1985.

J. C. Cox, S. A. Ross, and M. Rubinstein. Option pricing: A simplified approach. *Journal of financial Economics*, 7(3):229–263, 1979.

F. Delbaen and W. Schachermayer. A general version of the fundamental theorem of asset pricing. *Mathematische annalen*, 300(1):463–520, 1994.

K. Demeterfi, E. Derman, M. Kamal, and J. Zou. More than you ever wanted to know about volatility swaps. *Goldman Sachs Quantitative Strategies Research Notes*, 1999.

D. A. Dickey and W. A. Fuller. Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American statistical association*, 74(366a):427–431, 1979.

S. Dowling and J. Muthuswamy. The implied volatility of Australian index options. *Available at SSRN*, 2005. URL `http://ssrn.com/abstract=500165`.

B. Dupire. Model art. *Risk*, 6(9):118–124, 1993.

L. Ederington and W. Guan. Measuring historical volatility. *Journal of Applied Finance*, 16 (1), 2006.

L. H. Ederington and W. Guan. How asymmetric is U.S. stock market volatility? *Journal of Financial Markets*, 13(2):225–248, 2010.

W. Egan. The distribution of S&P 500 index returns. *Available at SSRN*, 2007. URL `http://ssrn.com/abstract=955639`.

E. F. Fama. The behavior of stock-market prices. *Journal of business*, pages 34–105, 1965.

S. Figlewski and X. Wang. Is the 'leverage effect' a leverage effect? *Available at SSRN*, 2000. URL `http://ssrn.com/abstract=256109`.

Stephen Figlewski. Forecasting volatility. *Financial Markets, Institutions & Instruments*, 6 (1):1–88, 1997.

J. Fleming, B. Ostdiek, and R. E.S Whaley. Predicting stock market volatility:A new measure. *Journal of Futures Markets*, 15(3):265–302, 1995.

B. Frijns, C. Tallau, and A. Tourani-Rad. The information content of implied volatility: evidence from Australia. *Journal of Futures Markets*, 30(2):134–155, 2010.

G. L. Gastineau. An index of listed option premiums. *Financial Analysts Journal*, pages 70–75, 1977.

P. Giot. Relationships between implied volatility indexes and stock index returns. *The Journal of Portfolio Management*, 31(3):92–100, 2005.

I. V. Girsanov. On transforming a certain class of stochastic processes by absolutely continuous substitution of measures. *Theory of Probability & Its Applications*, 5(3):285–301, 1960.

L. G. Godfrey. Testing against general autoregressive and moving average error models when the regressors include lagged dependent variables. *Econometrica: Journal of the Econometric Society*, pages 1293–1301, 1978.

D. Goldstein and N. Taleb. We don't quite know what we are talking about when we talk about volatility. *Journal of Portfolio Management*, 33(4), 2007.

M. T. González and A. Novales. Are volatility indices in international stock markets forward looking? *Revista de la Real Academia de Ciencias Exactas, Fisicas y Naturales. Serie A. Matematicas*, 103(2):339–352, 2009.

M. T. González and A. Novales. The information content in a volatility index for Spain. *SERIEs: Journal of the Spanish Economic Association*, 2(2):185–216, 2011.

C. W. J. Granger. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica: Journal of the Econometric Society*, pages 424–438, 1969.

W. Guo and M. E. Wohar. Identifying regime changes in market volatility. *Journal of Financial Research*, 29(1):79–93, 2006.

J. D. Hamilton. *Time series analysis*, volume 2. Cambridge University Press, 1994.

J. M. Harrison and S. R. Pliska. Martingales and stochastic integrals in the theory of continuous trading. *Stochastic processes and their applications*, 11(3):215–260, 1981.

C. R. Harvey and R. E. Whaley. Dividends and S&P 100 index option valuation. *Journal of Futures Markets*, 12(2):123–137, 1992.

E. G. Haug. *The complete guide to option pricing formulas*, volume 2. McGraw-Hill New York, 2007.

Z. He and K. Maekawa. On spurious Granger causality. *Economics Letters*, 73(3):307–313, 2001.

J. Hull. *Options, futures, and other derivatives*. Pearson, 2009.

K. Itō. On stochastic differential equations. *Memoirs of the American Mathematical Society*, 4:1–51, 1951.

C. M. Jarque and A. K. Bera. A test for normality of observations and regression residuals. *International Statistical Review*, pages 163–172, 1987.

G. Jiang and Y. Tian. The model-free implied volatility and its information content. *Review of Financial Studies*, 18(4):1305–1342, 2005.

G. Jiang and Y. Tian. Extracting model-free volatility from option prices: An examination of the VIX index. *Journal of Derivatives*, 14(3), 2007.

S. S. S. Kumar. A first look at the properties of India's volatility index. *International Journal of Emerging Markets*, 7(2):160–176, 2012.

D. Kwiatkowski, P. C. B. Phillips, P. Schmidt, and Y. Shin. Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? *Journal of Econometrics*, 54(1):159–178, 1992.

J. Liu, S. Wu, and J.V. Zidek. On segmented multivariate regression. *Statistica Sinica*, 7: 497–526, 1997.

A. M. Malz. Do implied volatilities provide early warnings of market stress? *Available at SSRN*, 2000. URL `http://ssrn.com/abstract=211448`.

B. Mandelbrot. The variation of certain speculative prices. *Journal of business*, pages 394–419, 1963.

R. C. Merton. Theory of rational option pricing. *The Bell Journal of Economics and Management Science*, pages 141–183, 1973.

S. N. Neftci. *An introduction to the mathematics of financial derivatives*. Academic Press, 2000.

Anthony Neuberger. The log contract: a new instrument to hedge volatility. *The Journal of Portfolio Management*, 20(2):74–80, 1994.

W. K. Newey and K. D. West. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica: Journal of the Econometric Society*, pages 703–708, 1987.

O. Nikodym. Sur une généralisation des intégrales de M. J. Radon. *Fundamenta Mathematicae*, 15(1):131–179, 1930.

B. Øksendal. *Stochastic differential equations: An introduction with applications*. Springer, 1998.

A. Peiró. The distribution of stock returns: International evidence. *Applied Financial Economics*, 4(6):431–439, 1994.

P. Perron. L'Estimation de modèles avec changements structurels multiples. *L'Actualité économique*, 73(1-2-3), 1997.

E. Platen and R. Rendek. Empirical evidence on student-t log-returns of diversified world stock indices. *Journal of statistical theory and practice*, 2(2):233–251, 2008.

S.H. Poon and C.W.J. Granger. Forecasting volatility in financial markets: A review. *Journal of Economic Literature*, 41(2):478–539, 2003.

S. A. Ross. Options and efficiency. *The Quarterly Journal of Economics*, 90(1):75–89, 1976.

F. D. Rouah and G. Vainberg. *Option pricing models and volatility using Excel-VBA*, volume 361. Wiley, 2007.

G. Schwarz. Estimating the dimension of a model. *The annals of statistics*, 6(2):461–464, 1978.

G. W. Schwert. Why does stock market volatility change over time? *The Journal of Finance*, 44(5), 1989.

G.W. Schwert. Stock volatility and the crash of'87. *Review of Financial Studies*, 3(1):77–102, 1990.

D. P. Simon. The Nasdaq volatility index during and after the bubble. *The Journal of Derivatives*, 11(2):9–24, 2003.

C. Siriopoulos and A. Fassas. Implied volatility indices – a review. *Available at SSRN*, 2009. URL `http://ssrn.com/abstract=1421202`.

G. Skiadopoulos. The Greek implied volatility index: construction and properties. *Applied Financial Economics*, 14(16):1187–1196, 2004.

N. N. Taleb. *Fooled by randomness: The hidden role of chance in life and in the markets*. London: Texere, 2001.

H. A. Ting. Fear in the Korea market. *Review of Futures Markets*, 16(1):106–140, 2007.

H. Y. Toda and T. Yamamoto. Statistical inference in vector autoregressions with possibly integrated processes. *Journal of Econometrics*, 66(1):225–250, 1995.

R. S. Tsay. *Analysis of financial time series*, volume 543. Wiley-Interscience, 2005.

R. E. Whaley. Derivatives on market volatility. *The journal of Derivatives*, 1(1):71–84, 1993.

R. E. Whaley. The investor fear gauge. *The Journal of Portfolio Management*, 26(3):12–17, 2000.

R. E. Whaley. Understanding the VIX. *The Journal of Portfolio Management*, 35(3):98–105, 2009.

Appendix A

A.1 The equivalence of model-free implied variance and the fair value of future variance

We establish the theoretival equivalence of model-free implied volatility and the fair value of future variance (FVFV) following Jiang and Tian (2007). To start with, consider the definition of FVFV in Demeterfi et al. (1999): It is the risk-neutral expected value of the average (integrated) realized variance over a future time period from $t = 0$ to $t = T$: $E[\text{Var}_T] = \frac{1}{T}\int_0^T \sigma_t^2 dt$. Using the Jiang and Tian (2005) MFIV as described in Equation (3.4), this can be written as

$$E[\text{Var}_T] = \frac{2}{T}\int_0^\infty \frac{e^{rT}C(T,K) - (S_0e^{rT} - K)^+}{K^2}dK. \tag{A.1}$$

Taking e^{rT} out of the integral and partitioning at $F_0 = S_0e^{rT}$ gives

$$\begin{aligned}
E[\text{Var}_T] &= \frac{2e^{rT}}{T}\left[\int_0^{S_0e^{rT}} \frac{C(T,K) - (S_0 - Ke^{-rT})^+}{K^2}dK\right. \\
&\quad \left. + \int_{S_0e^{rT}}^\infty \frac{C(T,K) - (S_0 - Ke^{-rT})^+}{K^2}dK\right]. \\
&= \frac{2e^{rT}}{T}\left[\int_0^{S_0e^{rT}} \frac{C(T,K) - (S_0 - Ke^{-rT})^+}{K^2}dK + \int_{S_0e^{rT}}^\infty \frac{C(T,K)}{K^2}dK\right].
\end{aligned} \tag{A.2}$$

The put-call parity[1] then implies that

$$
\begin{aligned}
&= \frac{2e^{rT}}{T}\left[\int_0^{S_0e^{rT}} \frac{P(T,K)}{K^2}dK + \int_{S_0e^{rT}}^{\infty} \frac{C(T,K)}{K^2}dK\right] \\
&= \frac{2e^{rT}}{T}\left[\int_0^{F_0} \frac{P(T,K)}{K^2}dK + \int_{F_0}^{\infty} \frac{C(T,K)}{K^2}dK\right]. \qquad \text{(A.3)}
\end{aligned}
$$

Equation (A.3) is comparable to the compact expression for the Jiang and Tian (2005) MFIV given in Equation (3.5). Following Jiang and Tian (2007), it can be rewritten as

$$
\begin{aligned}
E[\mathrm{Var}_T] &= \frac{2e^{rT}}{T}\left[\int_0^{S_*} \frac{P(T,K)}{K^2}dK + \int_{S_*}^{\infty} \frac{C(T,K)}{K^2}dK + \int_{S_*}^{F_0} \frac{P(T,K)-C(T,K)}{K^2}dK\right] \\
&= \frac{2e^{rT}}{T}\left[\int_0^{S_*} \frac{P(T,K)}{K^2}dK + \int_{S_*}^{\infty} \frac{C(T,K)}{K^2}dK + \int_{S_*}^{F_0} \frac{Ke^{-rT}-S_0}{K^2}dK\right],
\end{aligned}
\qquad \text{(A.4)}
$$

where $S_* < F_0$. The third term can be integrated, so that after some lengthy manipulations we arrive at an expression for the model-free variance which is exactly the same as the formulation of the fair value of future variance given by Demeterfi et al. (1999) in their equation (26):

$$
\begin{aligned}
E[\mathrm{Var}_T] &= \frac{2e^{rT}}{T}\left[\int_0^{S_*} \frac{P(T,K)}{K^2}dK + \int_{S_*}^{\infty} \frac{C(T,K)}{K^2}dK + \left(\ln(K)e^{-rT} + \frac{S_0}{K}\right)\Big|_{S_*}^{F_0}\right] \\
&= \frac{2}{T}\left[rT - \left(\frac{S_0}{S_*}e^{rT} - 1\right) - \ln\left(\frac{S_*}{S_0}\right)\right. \qquad \text{(A.5)} \\
&\quad \left. + \, e^{rT}\left(\int_0^{S_*} \frac{P(T,K)}{K^2}dK + \int_{S_*}^{\infty} \frac{C(T,K)}{K^2}dK\right)\right].
\end{aligned}
$$

[1] Recall that for a European option the put-call parity is given by $C + Ke^{-rT} = P + S_0$.

A.2 Portfolio replication strategy

We derive the static portfolio replication argument as proposed by Carr and Madan (1998). Recall that the argument states that any twice differentiable function $f(S)$ can be re-written as:

$$f(S) = f(\kappa) + f'(\kappa)(S-\kappa) + \int_0^{\kappa} f''(K)(K-S)^+ dK + \int_{\kappa}^{\infty} f''(K)(S-K)^+ dK. \quad \text{(A.6)}$$

To see this, first let $\delta(K)$ be an impulse symbol[2] characterized by

$$\delta(K) = \begin{cases} 0 & \text{if } K \neq 0 \\ +\infty & \text{if } K = 0 \end{cases} \quad \text{and} \quad \int_{-\infty}^{\infty} \delta(K) dK = 1. \quad \text{(A.7)}$$

Moreover, it is the derivative of the Heaviside step function $H(K)$ (see e.g. Bracewell 2000):

$$\delta(K) = H'(K), \quad H(K) = \begin{cases} 0 & \text{if } K < 0 \\ 1 & \text{if } K > 0 \end{cases} \equiv \mathbf{1}_{K>0}.$$

Note also that $H(K)$ is the derivative of $(K)^+ = \max(0, K)$. The impulse symbol has the following *sifting property*

$$\int_{-\infty}^{+\infty} f(K)\delta(K-a)dK = f(a). \quad \text{(A.8)}$$

This is because $\delta(K-a)$ is zero everywhere except at $K = a$, which allows for restricting the range of the integral to an epsilon interval around a, so that

$$\int_{-\infty}^{+\infty} f(K)\delta(K-a)dK = \int_{a-\epsilon}^{a+\epsilon} f(K)\delta(K-a)dK,$$

where $\epsilon > 0$ is an infinitesimal number. Inside this interval f is approximately constant and can thus be pulled out of the integral, meaning that

$$\int_{a-\epsilon}^{a+\epsilon} f(K)\delta(K-a)dK = \int_{a-\epsilon}^{a+\epsilon} f(a)\delta(K-a)dK = f(a)\int_{a-\epsilon}^{a+\epsilon} \delta(K-a)dK = f(a).$$

[2] Also known as the Dirac delta function after Paul Dirac, an English theoretical physicist, for his pioneering work on quantum mechanics. It can be considered a generalized function with infinite height, zero width and an area of one.

Let $f(S)$ be a twice-differentiable function defined in $\mathbb{R}$. The sifting property and properties of the integral now imply that

$$\begin{aligned} f(S) &= \int_0^\infty f(K)\delta(S-K)dK \\ &= \int_0^\kappa f(K)\delta(S-K)dK + \int_\kappa^\infty f(K)\delta(S-K)dK \\ &= \underbrace{\int_0^\kappa f(K)\delta(K-S)dK}_{I_1} + \underbrace{\int_\kappa^\infty f(K)\delta(S-K)dK}_{I_2} \end{aligned} \tag{A.9}$$

for some threshold $\kappa > 0$. Applying integration by parts[3] to I_1 with $u = f(K)$, $dv = \delta(K-S)$, $du = f'(K)$ and $v = \mathbf{1}_{K>S}$ gives

$$\begin{aligned} I_1 &= \int_0^\kappa f(K)\delta(K-S)dK \\ &= f(K)\mathbf{1}_{K>S}\Big|_0^\kappa - \int_0^\kappa f'(K)\mathbf{1}_{K>S}dK \\ &= f(\kappa)\mathbf{1}_{\kappa>S} - \int_0^\kappa f'(K)\mathbf{1}_{K>S}dK. \end{aligned} \tag{A.10}$$

For I_2, choose $u = f(K)$, $dv = \delta(S-K)$, $du = f'(K)$ and $v = -\mathbf{1}_{K>S}$, so that

$$\begin{aligned} I_2 &= \int_\kappa^\infty f(K)\delta(S-K)dK \\ &= -f(K)\mathbf{1}_{S>K}\Big|_\kappa^\infty + \int_\kappa^\infty f'(K)\mathbf{1}_{K<S}dK \\ &= f(\kappa)\mathbf{1}_{S>\kappa} + \int_\kappa^\infty f'(K)\mathbf{1}_{K<S}dK. \end{aligned} \tag{A.11}$$

Substitute I_1 and I_2 into equation (A.9) to obtain

$$f(S) = f(\kappa) - \underbrace{\int_0^\kappa f'(K)\mathbf{1}_{K>S}dK}_{I_3} + \underbrace{\int_\kappa^\infty f'(K)\mathbf{1}_{K<S}dK}_{I_4}. \tag{A.12}$$

Integrating by parts once again, we choose $u = f'(K)$, $v = \mathbf{1}_{K>S}$, $du = f''(K)$ and

[3] The conventional technique for integration by parts is given by $\int udv = uv - \int vdu$.

$dv = (K-S)^+$. Thus

$$\begin{aligned}
I_3 &= \int_0^{\kappa} f'(K)\mathbf{1}_{K>S} dK \\
&= f'(K)(K-S)^+\Big|_0^{\kappa} - \int_0^{\kappa} f''(K)(K-S)^+ dK \\
&= f'(\kappa)(\kappa-S)^+ - \int_0^{\kappa} f''(K)(K-S)^+ dK.
\end{aligned} \tag{A.13}$$

Finally, for I_4, choosing $u = f'(K)$, $v = \mathbf{1}_{S>K}$, $du = f''(K)$ and $dv = -(S-K)^+$ yields

$$\begin{aligned}
I_4 &= \int_{\kappa}^{\infty} f'(K)\mathbf{1}_{K<S} dK \\
&= -f'(K)(S-K)^+\Big|_{\kappa}^{\infty} + \int_{\kappa}^{\infty} f''(K)(S-K)^+ dK \\
&= -f'(\kappa)(S-\kappa)^+ + \int_{\kappa}^{\infty} f''(K)(S-K)^+ dK.
\end{aligned} \tag{A.14}$$

Substitute I_3 and I_4 back to equation (A.12) to obtain

$$\begin{aligned}
f(S) = f(\kappa) &+ f'(\kappa)\left[(\kappa-S)^+ - (S-\kappa)^+\right] \\
&+ \int_0^{\kappa} f''(K)(K-S)^+ dK + \int_{\kappa}^{\infty} f''(K)(S-K)^+ dK.
\end{aligned} \tag{A.15}$$

Since[4] $(\kappa-S)^+ - (S-\kappa)^+ = S-\kappa$, equation (A.15) can be written

$$f(S) = f(\kappa) + f'(\kappa)(S-\kappa) + \int_0^{\kappa} f''(K)(K-S)^+ dK + \int_{\kappa}^{\infty} f''(K)(S-K)^+ dK,$$

which is equation (A.6). As Carr and Madan (1998) point out, the first term is equal to the payoff of a static long position on $f(\kappa)$ bonds paying one currency unit at time T. The second term corresponds to the payoff from a long position on $f'(\kappa)$ calls and a short position on $f'(\kappa)$ puts, both with strike κ. The third term arises from $f''(K)dK$ puts at all strikes less than κ. Similarly, the fourth term represents a static position in an equal amount of calls at all strikes greater than κ.

[4]This can be easily verified by examining the cases $S > \kappa$ and $S < \kappa$ separately.