Anomalies of the Volatility Surface

Structural Errors and Speculative Bubbles

TECHNISCHE
UNIVERSITÄT
DARMSTADT

Philip Maximilian Stahl

Dissertation submitted in fulfilment of the requirements for the degree of
Doctor rerum politicarum (Dr. rer. pol.)

Supervisor	**Prof. Dr. Dirk Schiereck**
Second Advisor	**Prof. Dr. Dominik Wolff**

Department of Business Administration, Economics, and Law
Technische Universität Darmstadt
Germany

Darmstadt, 2024

Stahl, Philip: Anomalies of the Volatility Surface – Structural Errors and Speculative Bubbles
Darmstadt, Technische Universität Darmstadt,
Year thesis published in TUprints: 2024
Date of disputation: April 18th, 2024

To Marianne, Udo, and Jianan.

Abstract

This dissertation examines anomalies of the volatility surface. Technical anomalies as well as market anomalies are examined. The VIX Index structurally underestimates model-free implied variance, and the extend of underestimation varies over time. The SABR model can be utilized to identify stock price bubbles, and arbitrage opportunities in the volatility surface coincide with other changes in trading dynamics in the underlying.

Kurzzusammenfassung

Diese Dissertation untersucht Anomalien der Volatilitätsoberfläche. Es werden sowohl technische Anomalien als auch Marktanomalien untersucht. Der VIX-Index unterschätzt strukturell die modellfreie implizite Varianz, und das Ausmaß der Unterschätzung variiert im Laufe der Zeit. Das SABR-Modell kann zur Identifizierung von Aktienpreisblasen verwendet werden, und Arbitragemöglichkeiten auf der Volatilitätsoberfläche fallen mit anderen Änderungen in der Handelsdynamik des Basiswerts zusammen.

Contents

1 Introduction **1**

1.1 Background and Motivation 1

1.2 Research Questions 3

1.3 Structure of the Dissertation 5

2 Structural Errors in Variance Swap Replication **9**

2.1 Introduction 9

2.2 Theoretical Background and Fundamental Challenges 12

2.2.1 Derivation of Model-free Implied Variance 12

2.2.2 Discretization and Approaches to Interpolation 13

2.2.3 Truncation and Approaches to Extrapolation 15

2.3 Implementations 16

2.3.1 The CBOE VIX Index 16

2.3.2 Implementation by Carr and Wu (2009) 18

2.3.3 Implementation by Jiang and Tian (2007) 18

2.4 Experiment Design 19

2.4.1 Model-specific Forward Variance 20

2.4.2 Simulated Discretization 21

2.4.3 Simulated Truncation 22

2.5 Stochastic Models and Model-True Variance 23

2.5.1 The Heston Stochastic Volatility Model 23

2.5.2 The Merton Jump-Diffusion Model 25

2.5.3 The Stochastic Volatility plus Jumps (SVJ) Model . . 26

2.6 Analysis of Error Dynamics . . . 27
2.6.1 Error Dynamics of Stochastic Volatility . . . 27
2.6.2 Error Dynamics under Jumps in the Price Process . . . 30
2.6.3 Error Dynamics under Stochastic Volatility and Jumps 33
2.7 Summary of Findings . . . 34
2.8 Conclusion . . . 35

3 Asymptotic Extrapolation of Model-free Implied Variance 37
3.1 Introduction . . . 38
3.1.1 Notation . . . 40
3.1.2 Literature Review . . . 41
3.2 Structural Shortcomings of the VIX Index . . . 45
3.3 Asymptotic Extrapolation and Construction of the Truncation Error . . . 46
3.3.1 Construction of the Truncation Error . . . 46
3.3.2 Minimum Requirements for Strike Grid Width . . . 50
3.3.3 Convergence of TE_{Left} and TE_{Right} . . . 52
3.4 Applying the Truncation Error Term to the VIX Index . . . 53
3.4.1 Preliminary Modifications . . . 53
3.4.2 Step-by-Step Example . . . 55
3.5 Historical Analysis . . . 56
3.5.1 Dataset and Computation . . . 57
3.5.2 Analysis . . . 58
3.6 Conclusion . . . 64

4 Martingale Defects in the Volatility Surface and Bubble Conditions in the Underlying 67
4.1 Introduction . . . 68
4.2 Literature Review . . . 69
4.3 Theoretical Background . . . 73
4.4 Data and Calibration . . . 76
4.5 Analysis . . . 81
4.5.1 Defect Persistence . . . 82

4.5.2 Change of Distribution 83
4.5.3 Option Trading Activity 85
4.5.4 Short Interest Ratio 87
4.5.5 Institutional Ownership 89
4.5.6 Bubbles as Market-wide phenomenon 91
4.6 Conclusion 94

5 Concluding Remarks **97**

A Chapter 2 **113**
A.1 Calculation of the VIX Index 113
A.2 MFIV in terms of log-forward moneyness 114
A.3 Derivation of the extrapolated Black-Scholes call price 115
A.4 Dynamics of Lee [2004]'s tail parameter 116
A.5 Truncation error under different cutoff strikes 117

B Chapter 3 **119**
B.1 Calibration results 119
B.2 Total trading activity over time 119
B.3 Relative trading activity over time by maturity 120
B.4 Persistence 122
B.5 Results for Full Surface 122

List of Tables

3.1 CBOE-VIX Strike Contributions 55
3.2 Step-by-step Calculation Results 56

4.1 Aggregate results of two-sample Kolmogorov-Smirnov tests . . 84
4.2 Aggregate results for t-tests of the number of actively traded option contracts 85
4.3 Aggregate results for t-tests of outstanding short interest divided by ADTV 87
4.4 Aggregate results for t-tests of the percentage of institutional ownership 89

B.1 Descriptive statistics of calibration results 120
B.2 Tenor classifications 121
B.3 Number of events and number of unique affected companies for different persistence and threshold configurations 123
B.4 Aggregate results for t-tests of the number of actively traded option contracts 124
B.5 Aggregate results for t-tests of outstanding short interest divided by ADTV (Full surface) 124
B.6 Aggregate results for t-tests of the percentage of institutional ownership (Full surface) 125

List of Figures

1.1 Daily actively traded options contracts 2
1.2 Structure of the dissertation 6

2.1 Comparison of Implementation Schemes 16
2.2 Example Strike Price Grids 22
2.3 Comparison under Heston Model, Variation of v_0 28
2.4 Comparison under Heston Model, Variation of $\bar{v}$ 30
2.5 Comparison under Heston Model, Variation of κ 31
2.6 Comparison under Merton Model, Variation of α 32
2.7 Comparison under SVJ Model, Variation of η 34

3.1 Lee (2004)-based Extrapolation 50
3.2 Lee (2004) Tail Parameter Dynamics 51
3.3 Historical Truncation Error for the S&P 500 Index 59
3.4 Historical Truncation Error for the DJIA Index 61
3.5 Historical Truncation Error of the VIX Index 62

4.1 Analysis of persistence across thresholds 79
4.2 Intertemporal clustering of bubble events 92

A.1 Beta-parameter Simulation 116
A.2 Minimum Cutoff Strike Comparison 117

B.1 Number of contracts available used for calibration 121
B.2 Analysis of trading activity by tenor 122

List of Abbreviations

ADTV ..... Average Daily Trading Volume

ATM ..... At-the-money Option Strike

BSIV ..... Black-Scholes Implied Volatility

BSM Model ..... Black-Scholes-Merton Model

CBOE ..... Chicago Board of Options Exchange

CEV Model ..... Constant Elasticity of Variance Model

CIR-process ..... Cox-Ingersoll-Ross-process

DE ..... Discretization Error

DJIA Index ..... CBOE Dow Jones Industrial Average Index

DTE ..... Days to Expiration

G-SVJD Model ..... Generalized-SVJD Model

IV ..... Implied Volatility

K-S test ..... Kolmogorov-Smirnov Test

MFIV ..... Model-free Implied Variance

MSCI IMI ..... MSCI Investible Markets Index

MTV ..... Model-True Variance

NCS ..... Natural Cubic Splines

ND ..... No Dominance

NFLVR ..... No Free Lunch With Vanishing Risk

SABR Model ..... Stochastic Alpha, Beta, Rho Model

SDE ..... Stochastic Differential Equation

SPX Index ..... Standard & Poor's 500 Index

SVJ Model ..... Stochastic Volatility plus Jumps Model

SVJD Model ..... Stochastic Volatility Jump Diffusion Model

TE ..... Truncation Error

VIX Index ..... CBOE Volatility Index

VS ..... Variance Swap

VVIX Index ..... CBOE Volatility-of-VIX Index

Chapter 1

Introduction

1.1 Background and Motivation

> *When judged by its ability to explain the empirical data, option pricing theory is the most successful theory not only in finance, but in all of economics.* [Ross, 1987]

While option contracts have a long history, the breakthrough in option pricing were the seminal papers by Black and Scholes [1973] and Merton [1973] who independently developed a closed-form solution for European options in continuous time and introduced the market's expectation of future volatility as explicit input parameter of the pricing equation. This contribution revolutionized financial markets and garnered widespread recognition, ultimately earning Robert C. Merton and Myron S. Scholes the Nobel Memorial Prize in Economic Sciences for their transformative work[1]. The concept of implied volatility emerged as integral part of option pricing theory and as a widely used standard to quote option prices in business transactions. It is calculated by solving the *Black-Scholes-Merton model (BSM model)*[2] for its variance pa-

[1] Fisher Black passed away in 1995, and Nobel Prizes are not awarded posthumously.

[2] It is possible to calculate the implied volatility for other option pricing models, but

rameter given an option contract specification plus its observed price. The collection of all implied volatilities across strikes and expirations at a point in time is called the volatility surface, and it represents the market's collective expectation of future uncertainty.

Subsequent decades witnessed the development of extensions and improvements of the Black-Scholes-Merton model, such as stochastic volatility (e.g. Heston [1993]) and jumps in asset prices (e.g. Merton [1976]), to better reflect real-world dynamics. The advances in theoretical option pricing lead to increased popularity of options. As market participants' understanding of pricing and hedging grew, liquidity increased, and more complex strategies, such as straddles, strangles, and spreads, became common tools to manage risk and speculate on market movements. Figure 1.1 illustrates the development of option trading based on the number of daily actively traded options contracts on the *Standard & Poor's 500 Index (SPX Index)* as well as on the constituents of the *MSCI Investible Markets Index (MSCI IMI)*, which covers the US' large-, mid-, and small-cap stocks. In recent years, option trading even gained popularity among retail investors, which may be seen as a testament to the success and increasing approachability of financial theory.

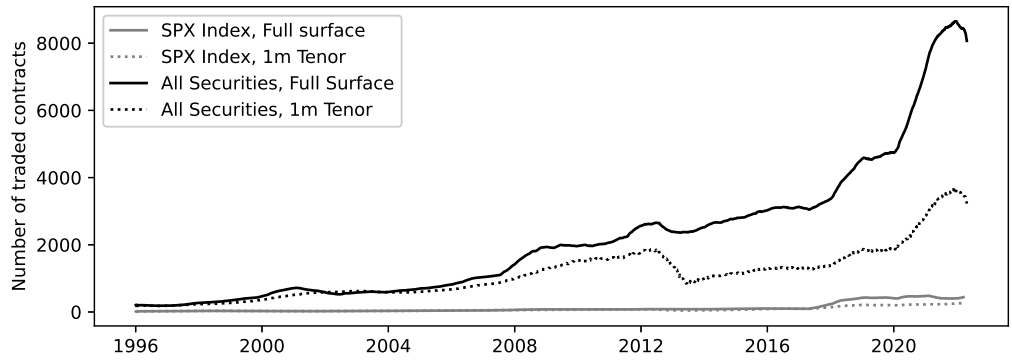

Figure 1.1: Daily actively traded options contracts for the SPX Index as well as the constituents of MSCI IMI. The contracts are split by tenor as defined in B.3. For visual clarity, one year rolling averages are presented.

Option pricing theory has become one of the most empirically accurate the-

rarely done in practice. We explore this issue in chapter 2.

ories in economics, and studying the volatility surface allows us to directly observe market perceptions of risk. It has been the subject of extensive study, with theoretical models and results that are fundamentally sound, and any known limitations can be effectively managed. However there are instances where the empirical volatility surface diverges from the theoretical predictions, even when the underlying theory stands true. This dissertation is concerned with these anomalies of the volatility surface.

This dissertation contributes to the understanding of the volatility surface in two perspectives and three distinct research questions. First, how technical reasons affect the analysis of aggregate measure of the future market uncertainty, and second, how irrational expectations and overoptimistic speculation can shape the volatility surface.

1.2 Research Questions

Within this dissertation, anomalies of the volatility surface are distinguished into two categories: *technical anomalies*, and *market anomalies*.

Technical anomalies arise due to practical limitations of option trading in the form of discrete and truncated strikes. They manifest as distinct gaps between strike prices where no options are specified or actively traded, and minimum and maximum strikes. In contrast, financial theory often relies on the foundational assumption of complete markets, wherein assets and derivatives are available at all possible strikes and without consideration for liquidity. This theoretical ideal posits a continuum of strike prices, ranging seamlessly from zero to infinity, necessitating interpolation to fill the gaps between strikes, and extrapolation to extend the observable data beyond the minimum and maximum strikes.

The first research question of this dissertation is to examine the effect of truncation and discretization on the estimation of the markets' expectation

of uncertainty in different market regimes. The analysis is based on simulated – and therefore known, not estimated – volatility surfaces, which are artificially truncated and discretized. From the remaining set of options, the implied volatility is calculated using different interpolation and extrapolation methods. The impact of discretization and truncation are examined and the ability of various suggested approaches are benchmarked. Variance swaps are volatility derivatives whose payoff is equal to the markets' expectation of future realized variance. By constructing a replication portfolio of an infinite continuum of out-of-the-money options, we are able to price a variance swap theoretically precise. Using a variety of option pricing models with different parameterizations to account for different market dynamics and regimes, we first generate infinitely dense and wide volatility surfaces. In separate tests, these surfaces are then artificially discretized from very dense to sparse, and artificially truncated from wide to narrow. A variance swap replication portfolio is then constructed from the remaining options, and the future realized variance is calculated using different approaches for inter- and extrapolation. For each model, we derive the model-true variance explicitly as a benchmark. In conjunction with the parameters of each simulated surface, this sets the benchmark against which the implied volatility is evaluated. We find that discretization bias in variance swaps tends to be self-correcting and the common approaches to interpolation are stable for reasonable strike grid densities. Truncation on the other hand introduces a structural negative bias, and none of the tested extrapolation methods provides a robust compensation. Depending on the specific method, the certain shapes of the volatility surface can lead to overcompensation, increasing the error.

The second research question focusses on this structural underestimation due to truncation further, to provide a reliable compensation term, and assess the effect of the truncation error in practice. Based on a well-established result on rational shapes of the volatility surface at extreme strikes by Lee [2004], we develop an explicit compensation term for the truncation-induced error. As variance swaps form the theoretical underpinning of volatility indices such as the CBOE VIX Index, the explicit compensation term permits us

to examine the effect of truncation on these indices. As the width of the actively traded option strike grid fluctuates over time, the severeness of the structural underestimation also fluctuates. In periods of positive returns with low volatility, the option strike grid contracts, leading to larger truncation errors. As prices correct and volatility increases, the option strike grid widens and the truncation error becomes negligible.

Market anomalies arise where the shape of the volatility surface diverges from fundamental financial theory, such that the fundamental theorem of asset pricing does not hold, and the subsequent theory breaks down. The third research question of this dissertation analyzes speculative bubbles in option prices, which manifest themselves in irrational shapes of the volatility surface. The *stochastic alpha beta rho model (SABR model)* is one of the few option pricing models that admits a strict martingale representation. We utilize this property by calibrating a SABR model to observed option prices to detect speculative bubbles. In an event study covering 2576 stocks over 26 years, we find that the option market absorbs speculative bubbles much faster than the equity market, and that irregular shapes in the volatility surface precede significant and persistent changes in the distribution of returns, option trading activity, outstanding short interest, and institutional ownership. We further find that speculative bubbles are temporally clustered, and occur almost exclusively in periods of sustained positive asset returns. During market corrections, speculative bubbles dissipate quickly.

1.3 Structure of the Dissertation

This dissertation consist of three separate studies. Figure 1.2 provides an overview of the structure of the dissertation and draws connections between the presented studies.

Chapter 2 examines the structural errors that arise when pricing variance

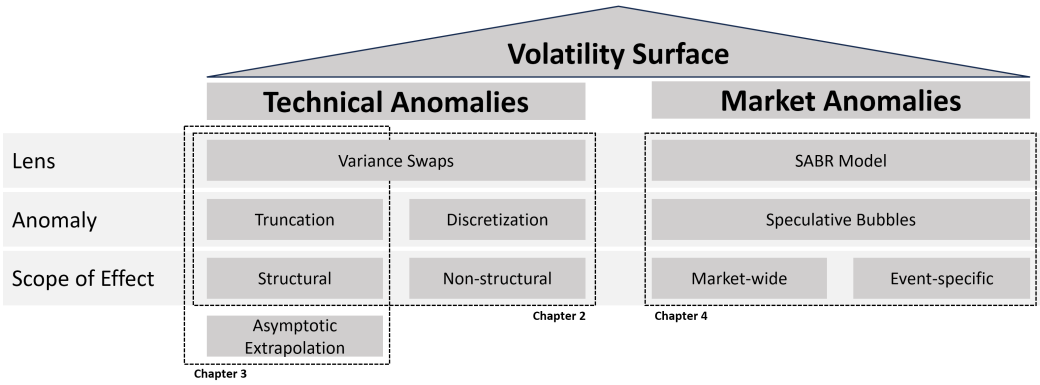

Figure 1.2: Structure of the dissertation. Anomalies in the volatility surface are distinguished into technical and market anomalies. Chapter 2 examines the technical anomalies in the volatility surface and the errors they introduce. Chapter 3 proposes a solution to the structural errors in the VIX Index which are caused by truncation. Chapter 4 studies the effect of irrational exuberance on the volatility surface.

swaps using replication portfolios based on artificially constrained option strike grids. It introduces the two sources of technical anomalies when pricing variance swaps, truncation and discretization, and examines their impact in detail [3].

Chapter 3 shows how the VIX Index is fundamentally a variance swap, and how it structurally underestimates model-free implied variance, and how this error depends on market volatility itself. It further suggests a novel approach to extrapolate the volatility smile in the strike domain using asymptotic properties derived from the underlying probability density function [4].

Chapter 4 analyzes irrational shapes in the volatility surface, and the implications of overoptimistic speculation. First, the notion of strict local martingales and their implications is introduced and the necessary conditions within the SABR model are examined. An event study then connects deformations

[3] Chapter 2 is currently under review in the *International Journal of Theoretical & Applied Finance*.

[4] Chapter 3 was published as Stahl, P.: Asymptotic extrapolation of model-free implied variance: exploring structural underestimation in the VIX Index. Review of Derivatives Research 25, 315–339 (2022) `https://doi.org/10.1007/s11147-022-09190-2`.

of the volatility smile with fundamental data and finds that it can indicate permanent changes in fundamental trading dynamics. The identified events are then studied in aggregate over time, showing that periods of sustained positive returns and low volatility increase the markets' optimism beyond a rational measure [5].

Finally, chapter 5 concludes the findings.

[5]Chapter 3 was published as Stahl, P. and Blauth, J.: Martingale Defects in the Volatility Surface and Bubble Conditions in the Underlying. Review of Derivatives Research 27, 85–111 (2024) https://doi.org/10.1007/s11147-023-09200-x.

Chapter 2

Structural Errors in Variance Swap Replication

Model-free implied variance, the theoretical underpinning of variance swaps as well as the VIX Index, is non-parametric by design. However, its implementation introduces some assumptions that cause structural errors. This chapter examines the performance of three commonly used implementations of model-free implied variance under a variety of simulated market conditions. We find that all three discussed implementations have drawbacks, and provide guidance for the choice of the implementation method.

2.1 Introduction

The CBOE VIX Index is the probably best known volatility index for equity markets. It attempts to extract the markets expectation of future realized variance of the S&P 500-Index. Under absence of arbitrage and risk neutrality, the fair price of any financial product equals its discounted expected payoff. For option contracts, the payoff is a function of the price of its under-

lying at expiration and the strike. Accordingly, the fair price is the payoff for each possible price of the underlying multiplied with its probability. Dupire [1994] shows that a complete option chain therefore fully determines the stochastic process of the underlying, and equivalently determines its risk-neutral density. Britten-Jones and Neuberger [2000] utilize this to extract the future realized variance of the underlying from its option chain without assuming a stochastic process for the underlying. This *model-free implied variance (MFIV)* is the theoretical foundation of the VIX Index as laid out in CBOE [2018].

In practice, option chains are naturally discrete, and any implementation of MFIV must consider two issues. First, the discreteness of option strikes requires interpolation. Second, the option chain has a minimum and a maximum strike, beyond which option prices are not available. While it is not strictly necessary to extrapolate option prices, many authors[1] chose to do so.

While MFIV itself is purposely mode-free, the choice of interpolation and extrapolation methods imply assumptions about the stochastic process of the underlying. The goal of this chapter is to show the effects of different interpolation and extrapolation approaches on MFIV under various market conditions. To accomplish this, three models are selected to benchmark against. They exhibit either stochastic variance, jumps in the return process, or both. For each model, the parameters of its primary feature are varied, and a multitude of volatility surfaces are generated. Each surface is artificially discretized and truncated, and three widely used implementations of MFIV are benchmarked against the models' expected variance.

We find that, due to the smooth nature of our simulated volatility surfaces, discretization appears to be of low concern. Naive linear interpolation approaches appear to be sufficient under a variety of market conditions. Truncation on the other hand poses a bigger challenge. Ignoring truncation, such as proposed by CBOE [2018], does lead to structural underestimation of

[1] An overview of related literature is provided in section 2.2.3; specifically see Jiang and Tian [2005], Jiang and Tian [2007], and Fengler [2009]

MFIV. Extrapolation with constant *Black-Scholes Implied Variance (BSIV)* also leads to underestimation of MFIV, albeit typically less. Spline-based extrapolation is prone to overshooting and unpredictability, as here the curvature of the variance smile at the cutoff point becomes a significant factor. Neither of these approaches fulfills no-arbitrage conditions.

This chapter provides the following contributions. We develop a systematic approach to analyze and compare the performance of MFIV implementations. Furthermore we identify the error dynamics under stochastic volatility and jump processes and the impact of their parameters on the performance of an implementation of MFIV. Most importantly, we find that all three implementations of MFIV exhibit structural errors that change with market dynamics. Any implementation of MFIV should be evaluated against a variety of stochastic processes for its underlying to achieve robustness in changing environments.

The remainder of this chapter is structured as follows. First, we derive MFIV and show the fundamental implementation challenges, and review the literature concerned with implementing MFIV. Second, we explain the three most common approaches to both issues in detail. Third, we set up the experiment design and show how simulated surfaces are discretized and truncated. Fourth, we derive the benchmark variance for each model. Fifth, we provide a detailed review of simulation results and the effects of market conditions on MFIV. Finally we summarize our findings and provide recommendations.

2.2 Theoretical Background and Fundamental Challenges

2.2.1 Derivation of Model-free Implied Variance

Under risk neutrality, option prices are determined by probability of the underlying expiring in-the-money and its respective payoff. Demeterfi et al. [1999] reverse-construct a trinomial strike-price tree with deterministic local volatility as a function of strike and time from observed option prices. Neuberger [1994] introduces the concept of futures contracts that pay the natural logarithm of an underlying future at expiration and shows that the payoff of a delta-hedged log-contract depends purely on volatility of the underlying, provided its variance is constant over time. Britten-Jones and Neuberger [2000] show that, with K being the option strike price, a $\frac{1}{K^2}$-weighted portfolio of out-of-the-money options has constant sensitivity to changes in variance, and construct a portfolio that perfectly replicates forward variance of a continuous underlying[2]. Furthermore, it is only based on the assumption of risk-neutrality, and does not rely on a specific stochastic model for the underlying process, except that it follows a diffusion. It is therefore called *model-free implied variance (MFIV)*. Britten-Jones and Neuberger [2000] define the MFIV as

$$\mathbb{E}_0^{\mathbb{Q}}\left[\int_0^T \left(\frac{dF_t}{F_t}\right)^2\right] = 2\int_0^\infty \frac{C(T,K) - max(F_0 - K, 0)}{K^2} dK \tag{2.1}$$

with $C(T,K)$ being the call option price, $K \in \boldsymbol{K}$ being the strike, and, since the expectation is taken under $\mathbb{Q}$, $C(t,K) = \mathbb{E}_t^{\mathbb{Q}}[max(F_t - K, 0)]$ for all $t \in \boldsymbol{T}$.

Since option sensitivity to changes in variance – commonly referred to as vega – peaks at the strike of the option, the discrete nature of strikes of observable options fundamentally introduces replication errors. Derman and

[2]For example, see Derman and Miller [2016], Chapter 4.

Miller [2016] show that the replication error grows larger as fewer contracts are observed. When few options are observed, the replication portfolio described by equation 3.2 exhibits fluctuating variance sensitivity in the put side, and diminishing variance in the call wing. This asymmetric sensitivity is a consequence of the $\frac{1}{K^2}$ weighting scheme, which compensates the higher dollar-vega of contracts with higher strikes.

Jiang and Tian [2005] show that replication portfolio holds for price processes with jumps in the price level. Furthermore, they find evidence that the MFIV is a more accurate predictor of future realized variance than historical variance or Black-Scholes ATM implied variance. They state theoretical upper bounds for the truncation and discretization errors. Jiang and Tian [2007] examine the implementation equation 3.2 by CBOE [2018] and find it to be systematically flawed. They suggest a numerical scheme to overcome the issues. This chapter extends their analysis and provides a framework to evaluate numerical schemes in the context of MFIV.

Carr et al. [2012] show that equation 3.2 also holds for arbitrary exponential Lévy processes with few technical conditions. Carr and Wu [2009] utilize MFIV to estimate variance risk premia for indices and single stocks. They also provide a detailed interpolation and extrapolation procedure that we will discuss in section 2.3.2.

2.2.2 Discretization and Approaches to Interpolation

Discretization errors result from the simple fact that option strikes are discrete by nature. Thus, we can only approximate the k-continuous integral of equation 3.2.

In their implementation of the VIX index, CBOE [2018] uses constant midpoint interpolation between observations, which we will discuss in section 2.3.1.

Carr and Wu [2009] interpolate implied variance linearly in similar manner and generate a high-density strike-price grid to integrate along. We will discuss their approach in more detail in 2.3.2. Kahale [2005] suggests interpolating observed call prices with polynomial splines, then translating it into Black-Scholes implied volatilities, and finally interpolating implied variances linearly. Fengler [2009] argues that this approach, to guarantee absence of arbitrage, requires absence of arbitrage within the observed option prices. Since this is not guaranteed, he suggests smoothing the volatility surface using natural cubic splines (NCS). When smoothing, true observations are removed in favour of an arbitrage-free surface calibration. Therefore, this type of interpolation is not a suitable approach in the context of MFIV computation. Laurini [2011] suggests a similar smoothing procedure, and impose absence of static arbitrage by using constrained B-splines. Fengler and Hin [2015] refine this by using a tensor-product B-spline to enforce no-arbitrage constraints in strike and calendar dimensions.

Jiang and Tian [2007] also chose natural cubic splines, but apply them directly to the implied volatilites of their observations. Their implementation will be discussed in 2.3.3.

Glau et al. [2017] suggest bivariate interpolation along the implied volatility smile based on Chebyshev polynomials. They find that their implementation approaches the accuracy of natural cubic splines at reduced computational expense. However, they do not apply their results to the specific problem at hand. Fukasawa et al. [2011] avoid numerical integration by integrating with respect to the standard normal density, and interpolating along the integrand using piecewise polynomials. They show that their approach prevents numerical oscillations of the error term. Some insight into these oscillations can also be found in Jiang and Tian [2005]. Loucks [1996] remarks that – in the context of terrain elevation – Chebyshev polynomials should only be used on fine grids, since they tend to overshoot when observations have large vertical distances. The same applies to volatility surface interpolation, and the level of trading activity of the chosen market should be considered when

choosing an interpolation method.

2.2.3 Truncation and Approaches to Extrapolation

Option strikes are naturally truncated. The defining integral of equation 3.2 requires option strikes from zero to infinity, which is naturally not observable.

The VIX Index implementation by CBOE [2018] does not extrapolate option prices beyond the outermost strikes. Accordingly, this leads to a structural underestimation of true MFIV, which depends on the distance of the outermost strikes and the relative price levels.

Another approach is to extrapolate option prices from the observed prices outward. As out-of-the-money option prices tend to approach zero as $|k| \to \infty$, it is standard practice to extrapolate on the BSIV surface[3]. Carr and Wu [2009] convert the outermost option prices into BSIV, and extrapolate with constant BSIV.

Jiang and Tian [2007] utilize the slope of the cubic spline at the outermost strikes, and extrapolate BSIV linearly with a constant slope. Fengler [2009] follows a similar approach on a smoothed surface.

A fundamentally different approach would be to calibrate a stochastic model against observable data and generate prices from the surface. This approach depends on wether the chosen model can accurately resemble the observed prices. In this chapter, we choose the stochastic model to compare different implementations, therefore a model-dependent extrapolation approach cannot be sensibly compared and will be omitted here. Figlewski [2008] extrapolates the tails by fitting a generalized extreme value distribution to option prices directly. Since this fundamentally implies the assumption of

[3]for example, see Andersen and Andreasen [2000], Cont and da Fonseca [2002], and Choi et al. [2009]

specific process dynamics, this approach is also not reviewed further.

2.3 Implementations

This section provides details on three implementations of equation 3.2. In the first section, the official definition of the VIX Index is examined. The following two sections provide different approaches to reduce the effect of discrete option prices on the computation of the continuous integral.

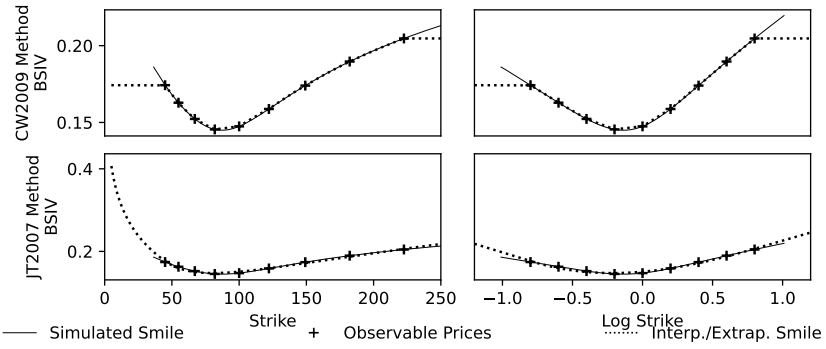

Figure 2.1: Comparison of three implementation schemes of equation 3.2. For visual clarity, out-of-the-money option prices have been translated into BSIV, even though the computation requires the corresponding prices. Option prices have been generated using a Heston model with parameters $\Omega = \{v_0 = 0.0225, \bar{v} = 0.0225, a = 3, \eta = 0.25, r = 0, \rho = 0.2\}$. Nine equidistant strikes centered around K_0 have been selected to be made available to the implementation function. The panels on the left side illustrate the implementation in dollar-strike terms, the panels on the right show the same in log-moneyness terms.

2.3.1 The CBOE VIX Index

The VIX Index is defined by CBOE [2018]. Their implementation focuses on computing a constant maturity volatility index of the S&P 500 Index.

For a chosen maturity from a given date, the observed option set is subjected to the following liquidity requirements.

1. From the complete set of options available for trading on a given day, select those with the chose maturity;

2. Choose all options that are currently out of the money;

3. Remove contracts which do not have valid bids;

4. If two strikes without bids are observed in succession, remove all further contracts.

The mid-price is calculated as the midpoint between bid and ask price for each contract.

The main VIX formula is based on a weighted Riemann sum which approximates equation 3.2 in the form

$$\text{VIX} = \sigma^2 \tag{2.2}$$

$$= \frac{2}{T}\Sigma_i \frac{\Delta K_i}{K_i^2} e^{rT} q(K_i) - \frac{1}{T}\left[\frac{F}{K^*} - 1\right]^2 \tag{2.3}$$

where K_i are the USD-Strikes of all $i \in [0, N]$ contracts, and q_i is the price of the respective out-of-the-money option at K_i. F is the forward price of the underlying, and K^* is the largest available strike below F. The intervals are defined as half the distance between the neigbouring strikes $\Delta K_i = \frac{K_{i+1} - K_{i-1}}{2}$. For the outermost strikes K_0 and K_N, this changes to $\Delta K_0 = K_1 - K_0$ and $\Delta K_N = K_N - K_{N-1}$. r is the risk-free rate, T is the remaining lifetime in

years[4].

This implementation implies that the interpolation between strikes assumes q to be constant around K_i, depending on the distance to its neighbours. While there is no specific term for extrapolation, the implementation via equation 2.2 implies an overhang with size $\frac{\Delta K_{0,N}}{2}$ for each tail.

2.3.2 Implementation by Carr and Wu (2009)

Carr and Wu [2009] propose using the difference between MFIV and ex-post realized variance as measure of the variance risk premium. Their analysis includes 5 stock indices and 35 individual stocks. As their diverse set of underlyings exhibits large differences in quote availability, they propose interpolating and extrapolating the strike-price grid before computing equation 3.2. In this chapter, our focus is examining their proposed implementation.

As out-of-the-money option prices approach zero as $|k| \to \pm\infty$, prices are first converted into BSIVs, inter- and extrapolation is applied, and the result is converted back into prices[5].

For a given set of implied volatilities, they suggest interpolating linearly between log-strikes. Outside of the observable strike range, implied volatility is assumed to be constant. Figure 2.1 illustrates this implementation in the top panels. This general approach – linear interpolation, flat extrapolation – has also been suggested by other authors, for example Jiang and Tian [2005].

From this interpolated and extrapolated volatility smile, $N_k = 2000$ equidistant log-strikes are drawn such that $k_{max} = -k_{min} = 8\sigma^2_{\text{Average BSIV}}$ with $\sigma^2_{\text{Average BSIV}}$ being the average BSIV of the entire set. With constant extrapolation, this very wide grid is simple to implement, but testing this method

[4]These definitions are exactly as stated in the implementation whitepaper CBOE [2018].

[5]This is standard practice, see e.g. Bates [1996], Bates [2000], and Jiang and Tian [2005].

against the surface of a parameterized stochastic process presents numerical challenges, unless implied volatilites are generated directly.

2.3.3 Implementation by Jiang and Tian (2007)

Jiang and Tian [2007] find that the implementation of equation 3.2 by CBOE [2018] leads to structural errors. They test this by simulating option prices from a *stochastic volatility plus jumps (SVJ)* model, and benchmarking equation 2.2 against the known model volatility. They show that the result of equation 2.2 may underestimate the true value by up to 198 basis points, or overestimate it by up to 79 basis points in their sample period from 1994 – 2004. They focus their simulations on a single volatility smile, and assume its maturity matches the desired horizon exactly.

To overcome truncation and discretization errors, they suggest interpolating and extrapolating option prices, and computing equation 2.2 from a subsample of these generated prices. As usual, interpolation and extrapolation are applied to the implied volatilities instead of prices, and are converted back into prices afterwards. To interpolate, Jiang and Tian [2007] suggest fitting *natural cubic splines (NCS)* to implied volatilites to generate a smooth volatility smile function that fits every single observed price exactly. They comment that a constant extrapolation such as chosen by Carr and Wu [2009] or Jiang and Tian [2005] violates no-arbitrage conditions, and instead suggest linearly extrapolating implied volatilities using the steepness of the NCS at the two outermost observations. Figure 2.1 illustrates this implementation in the bottom panels. It should be noted that this extrapolation is only linear in log-strike space (bottom right panel), and creates a very steep smile in dollar-strike space (bottom left panel).

Since they do not provide this information, we assume that $N_k = 2000$ equidistant log-strikes are generated, with $k_{max} = -k_{min} = 3\sigma^2_{\text{ATM BSIV}}$, with $\sigma^2_{\text{ATM BSIV}}$ being the at-the-money BSIV of the observed/simulated volatility

smile.

2.4 Experiment Design

To analyze the effect of variance surface dynamics on the pricing of variance swaps using numerical approaches, we specify a variety of stochastic processes as drivers for our underlying and compute their respective forward variances for a variety of parameters. Three different stochastic models are chosen to exhibit stochastic variance and jumps in isolation as well as in combination. We then generate decreasingly discretized and truncated sets of option prices, upon which we apply the three numerical approaches described in section 2.3. Finally, we benchmark the numerically computed MFIV against the analytically derived MTV as the strike grid width and density is increased.

2.4.1 Model-specific Forward Variance

To benchmark a numerical implementation, we need to compute the forward variance of the underlyings stochastic process. Since our analysis is based on option price surfaces of parameterized stochastic models, the process is known exactly, and we can compute the model-specific forward variance or *model-true variance (MTV)*. Alternatively, Jiang and Tian [2007] propose a Monte-Carlo approach.

The differences between different types of variance are crucial. The underlying's return process is governed by a parameterized stochastic process, and since the process and all parameters are known beforehand (in the simulation), we can take expectations and calculate MTV. At time $t = 0$, option prices are observed. Based on these observed option prices, a hedging portfolio is constructed via equation 3.2, and the ex-ante MFIV is calculated. At expiration at time $t = T$, the return process exhibits observable ex-post

historic variance. The stochastic process may or may not have a parameter referred to as *variance* or σ, but they bear different interpretations, and may exhibit very different values. The special case is the Black-Scholes model, where the stochastic parameter σ describes the variance of the underlyings' return process in its entirety, such that it is equivalent to the Black-Scholes MTV. As in this case the volatility smile is flat, and the MFIV is fundamentally a weighted average of the smile, MFIV is identical. In the Heston-model, as defined by the set of *stochastic differential equations (SDE)* 2.4, variance of the diffusion process is stochastic itself. In the Merton-model, as defined by SDE 2.13, the variance parameter σ only captures the variance of the diffusion process, and is only a part of the expected variance of the return process.

We will use the definition of a variance swap as MTV. Carr and Wu [2006] show that the fair price of a variance swap equals the average of the expected realized variance of underlying S_t for $t = 0 \to T$. Define

$$VS = \frac{\mathbb{E}_0^{\mathbb{Q}}[\int_{s=0}^{T} V_s ds]}{T}$$

as the fair price of this variance swap contract. Carr and Wu [2006] also show that, by definition, $VIX^2 = VS$.

To benchmark the numerical performance of the different implementations, we will analyze how each estimator approaches the MTV of a variety of parameterized stochastic processes. Model-specific MTV are described in detail in section 2.5.

2.4.2 Simulated Discretization

Strikes are chosen such that $k_{min} = -k_{max}$, and a center strike always exists at $k = 0$. Each grid is symmetric in log-moneyness and the strikes are equidistant. From each grid to the next, the number of strikes is increased

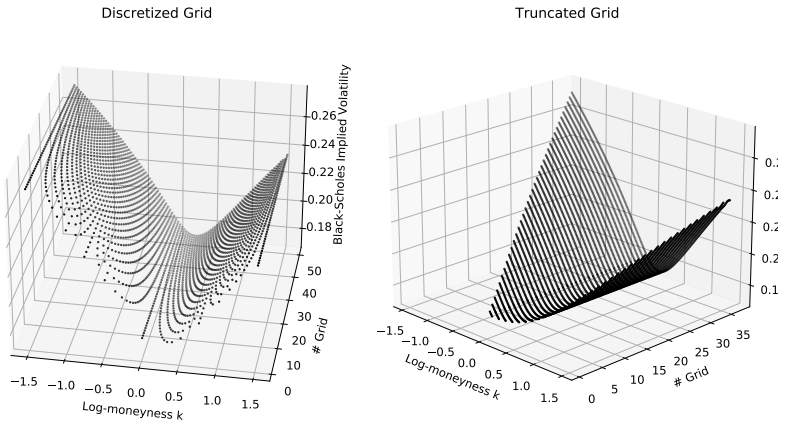

Figure 2.2: Example strike price grids, illustrating the generation of strike grids to test convergence. SVJ model with parameter set $\Omega_{\text{SVJ}} = \{v_0 = 0.0225, \bar{v} = 0.0225, \kappa = 3, \xi = 0.25, \rho = 0, \sigma = 0.22, z_1 = 0.51, z_2 = 0.48, \lambda = 0.225\}$ and time-to-maturity of one year.
Left panel shows discretized grids. 50 different grids, beginning with 5 strikes, up to 103 strikes. Between each grid, the number of strikes is increased by 2.
Right panel shows truncated grids. 50 different grids, symmetric in log-moneyness, beginning with $k_{max} = 0.02$, up to $k_{max} = 2$. Between grids, k_{max} is increased by 0.04.

by two. The first grid $g = 1$ is the most discrete, with five strikes in total such that for the first grid $\forall k^{g=1} \in \{-2, -1, 0, 1, 2\}$. The second grid $g = 2$ is less discrete, such that $\forall k^{g=2} \in \{-2, -1.33, -0.67, 0, 0.67, 1.33, 2\}$, and so forth for the remaining grids. The last grid has 103 equidistant strikes.

The left panel of figure 2.2 illustrates this using a parameterized SVJ model.

2.4.3 Simulated Truncation

To analyze the effect of truncation of the strike grid, a set of strike grids is generated, ranging from highly truncated to very wide. To isolate the effect from discretization, the strikes inbetween are chosen to be equidistant and densely spaced. In total, fifty strike grids are generated in log-moneyness space. Each grid is symmetrical in log-moneyness, such that $k_{min} = -k_{max}$. Strikes are chosen to be equidistant with a distance $k_{N+1} - k_N = 0.01$. The first grid $g = 1$ is the most truncated, with $k_{max} = 0.02$. For each subsequent grid, k_{max} is increased by 0.04, until $k_{max} = 2$.

The right panel of figure 2.2 illustrates this using a parameterized SVJ model.

The price of out-of-the-money options approaches 0 the further out the respective strike lies, which leads to numerical instability when solving the Black-Scholes equation for σ. This creates fluctuations in the volatility smile, making extrapolation even less reliable. The exact fluctuation depends on the stochastic model, the specific implementation of the root finder algorithm, and the computer itself. In our experiments, most volatility surfaces started to degrade beyond $k = \pm 2$, which is why we chose to limit all simulated surfaces to this range.

2.5 Stochastic Models and Model-True Variance

MFIV replicates the forward variance of the underlying, and the objective of this chapter is to understand the impact of truncation and discretization on the computation of MFIV. Since we choose specific stochastic processes to simulate the underlying, we can analytically derive its expected forward variance, which we will refer to as *model-true variance (MTV)*.

Since we are interested in the variance, all drift terms are assumed to be zero, and omitted in the SDEs.

2.5.1 The Heston Stochastic Volatility Model

In the Heston stochastic volatility model, stock returns follow a Brownian motion with stochastic instantaneous variance v_t, which is driven by a *Cox-Ingersoll-Ross (CIR)*-process, such that

$$dS_t = \sqrt{v_t} S_t (\rho dW_t^S + \sqrt{1-\rho^2} dW_t^v) \tag{2.4}$$

$$dv_t = \kappa(\bar{v} - v_t)dt + \xi\sqrt{v_t} dW_t^v \tag{2.5}$$

W^S and W^v are independent Wiener processes with constant instantaneous correlation ρ. The CIR process is mean reverting to long-term mean variance $\bar{v}$ with pullback speed κ.

It is important to note that, for longer dated options, the expected forward variance under $\mathbb{Q}$ does not depend on ξ. The forward variance is the sum of the variances of both SDEs. To show the independence from ξ, consider that, asymptotically[6], the probability distribution of v_t approaches the Gamma distribution with probability density function

$$v_\infty(\kappa, \bar{v}, \xi) = \frac{\beta^\alpha}{\Gamma(\alpha} \tag{2.6}$$

with $\Gamma(\alpha)$ being the Gamma function, and

$$\alpha = \frac{2\kappa}{\xi^2}$$

[6]The time scale of convergence depends on κ, typically in the order of a few days. In this analysis the focus lies on options with longer time to expiration, since short dated options often capture idiosyncratic effects.

$$\beta = \frac{2\kappa\bar{v}}{\xi}$$

By applying the Fourier transform, we arrive at the characteristic function of the long-run instantaneous variance as

$$\phi_v^{Heston}(u) = \left(1 - \frac{iu}{\beta}\right)^{-\alpha} \tag{2.7}$$

To compute the first moment of v_t, following Resnik [1998], take the partial derivative w.r.t the Fourier variable, divide by i, and evaluate at $u = 0$:

$$\mathbb{E}[v_t] = i^{-1} \frac{\partial \phi_v(u)}{\partial u}\bigg|_{u=0} \tag{2.8}$$

$$= i^{-1} \frac{i\alpha \left(1 - \frac{iu}{\beta}\right)^{-(1+\alpha)}}{\beta}\bigg|_{u=0} \tag{2.9}$$

Evaluation along the real line and simplifying then confirms the intuition that, for a mean reverting variance process, the expected long term variance exactly equals the long term variance level $\bar{v}$ and is independent from ξ.

$$\mathbb{E}[v_t] = \bar{v} \tag{2.10}$$

Broadie and Jain [2008a] present the fair strike of a variance swap on the Heston model as the sum of the variance of both SDEs.

$$MTV_{Heston} \coloneqq \mathbb{E}_0^{\mathbb{Q}}[S^2] = \mathbb{E}_0^{\mathbb{Q}}[\frac{1}{T}\int_0^T v_s ds] \tag{2.11}$$

$$= \frac{1 - e^{-\kappa T}}{\kappa T}(v_0 - \bar{v}) + \bar{v} \tag{2.12}$$

2.5.2 The Merton Jump-Diffusion Model

The Merton jump-diffusion model describes a stock price evolution as a continuous diffusion process with discrete jumps of random size and at random times. The diffusion component is identical to the Black-Scholes model. The jumps occur following a Poisson process P with jump intensity λ, and the jump size is lognormal distributed, with mean jump size z_1 and jump variance z_2. The jumps are persistent, such that the jump component is the sum of all jumps that happened up to time t. Therefore, $J_t = \Sigma_{i=1}^{P_t}[Y_i - 1]$ with $P_t \sim \text{Poisson}(\lambda)$ and $J_i \sim LN(z_1, z_2)$. Under $\mathbb{Q}$, the process is specified as

$$dS_t = \sigma dW_t^{\mathbb{Q}} S_t + \lambda(e^{z_1 + \frac{1}{2}z_2} - 1)S_t dt + S_t dJ_t \tag{2.13}$$

Broadie and Jain [2008b] analyze the effect of sampling on the fair price of variance swaps. We follow their approach to derive the continuous variance. At time T, P_T jumps have occurred, such that the realized variance of the jump component is $\frac{1}{T}\Sigma_{i=1}^{P_T}[\ln(Y_i)^2]$. The risk-neutral expectation of future total realized variance of S at time $t = 0$ is then given by

$$MTV_{Merton} \coloneqq \mathbb{E}_0^{\mathbb{Q}}[S^2] = \frac{1}{T}\int_0^T \sigma^2 dt + \frac{1}{T}\Sigma_{i=1}^{P_T}[\ln(Y_i)^2] \tag{2.14}$$

$$= \sigma^2 + \lambda(z_1^2 + z_2) \tag{2.15}$$

2.5.3 The Stochastic Volatility plus Jumps (SVJ) Model

To capture the stochastic nature of volatility, but also account for the high skewness in options with short remaining time to expiry, Bakshi et al. [2000] suggest combining the stochastic volatility dynamics of the Heston model with the jump-diffusion dynamics of the Merton model. The dynamics of

this SVJ model are described by

$$\frac{dS_t}{S_t} = \sqrt{v_t}(\rho dW_t^S + \sqrt{1-\rho^2}dW_t^v) + \sigma dW_t^{\mathbb{Q}} + \lambda(e^{z_1+\frac{1}{2}z_2^2} - 1)dt + dJ_t \tag{2.16}$$

$$dv_t = \kappa(\bar{v} - v_t)dt + \xi\sqrt{v_t}dW_t^v \tag{2.17}$$

with all parameters as described above. The components of the stochastic evolution are independent. Following Broadie and Jain [2008b], the forward variance is the sum of the individual component processes, such that

$$MTV_{SVJ} \coloneqq \mathbb{E}_0^{\mathbb{Q}}[S^2] = \frac{1-e^{-\kappa T}}{\kappa T}(v_0 - \bar{v}) + \bar{v} + \lambda(z_1^2 + z_2^2) \tag{2.18}$$

.

2.6 Analysis of Error Dynamics

Initial model parameters have been generated by fitting a model to observed data, and have been simplified and selected to illustrate the various dynamics of interest. For each simulation, a single parameter is increased and decreased within plausible constraints to generate a similar but sufficiently different smile.

Plots within the text have been limited to present only the most insightful results. Further results can be found in the appendix.

2.6.1 Error Dynamics of Stochastic Volatility

The Heston model tends to exhibit log-moneyness symmetrically shaped volatility smiles with upward sloping, but flattening tails. This behaviour is consistent with the mean-reverting nature of the variance process.

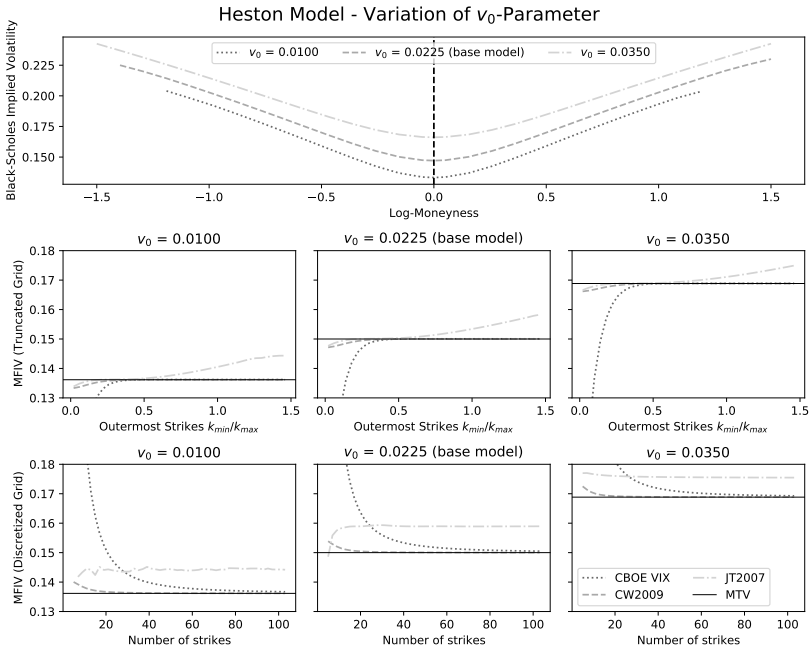

Figure 2.3: Comparison of implementation performance. Option prices have been generated from a Heston model with three parameter sets $\Omega = \{v_0 = [0.0100, 0.0225, 0.0350], \bar{v} = 0.0225, a = 3, \eta = 0.25, r = 0, \rho = 0\}$. Top panel shows implied volatility smile of each parameterization. Middle panels show computed MFIVs for each parameter set after truncation. Bottom panels show computed MFIVs for each parameter set after discretization.

Figure 2.3 shows the effect of truncation and discretization as described in section 2.4. The base model calibration is $\Omega = \{v_0 = 0.0225, \bar{v} = 0.0225, a = 3, \eta = 0.25, r = 0, \rho = 0\}$. One surface is generated for the base model, a second surface is generated with initial variance $v_0 = 0.0100$, and a third surface is generated with $v_0 = 0.0350$. All three surfaces are shown in the top panel. Only initial variance v_0 is varied.

In the middle row, for each calibration, the effect of truncation on the three MFIV implementations are shown. The middle left panel shows the error for the $v_0 = 0.0100$ calibration, the middle center panel shows this for the base model, and the middle right panel show this for the $v_0 = 0.0350$ calibration.

For every calibration, the respective MTV is shown as benchmark. The three bottom panels show the effect of discretization on the three implementations in identical order.

Comparing the three parameterizations it can be seen that varying v_0 impacts primarily the level of MTV, while the shape of the volatility smiles remains similar. The error dynamics are therefore very similar as well. The VIX Index implementation does not extrapolate, which leads to a strong underestimation of MTV when applied to a truncated surface. As the surface gets wider, this error disappears completely. The implementation by Carr and Wu [2009] extrapolates with constant variance, and compensates truncation partially. The implementation by Jiang and Tian [2007] extrapolates linearly using the slope of cubic splines, which compensates at first, but overcompensates quickly as the curvature of the smile changes at the truncation points. On a discretized option strike grid, all three implementations overestimate MTV. The VIX Index implementation interpolates with constant volatility and overestimates MTV by a large amount. The implementation by Carr and Wu [2009] and Jiang and Tian [2007] both interpolate using cubic splines, but differ significantly in their error dynamics. While the implementation by Carr and Wu [2009] approaches MTV quickly, the implementation by Jiang and Tian [2007] quickly approaches a constant volatility above MTV. This overestimation is caused by the extrapolation, as all discretized surfaces have identical strike grid width. For $v_0 = 0.0100$, we observe an irregular fluctuation which is caused by numerical errors when calculating option prices very far from the money.

Figure 2.4 shows similar behaviour when varying the long-term variance level $\bar{v}$. This results in volatility smiles with lower and wider spread ATM BSIV, and steeper wings. Under discretization, the estimation error dynamics behave very similar to above. The increased steepness and convexity in the wings affect the dynamics of the estimation error under truncation systematically. As extrapolation is missing in the implementation by CBOE [2018], its error dynamics are predictable in that they underestimate MTV severely

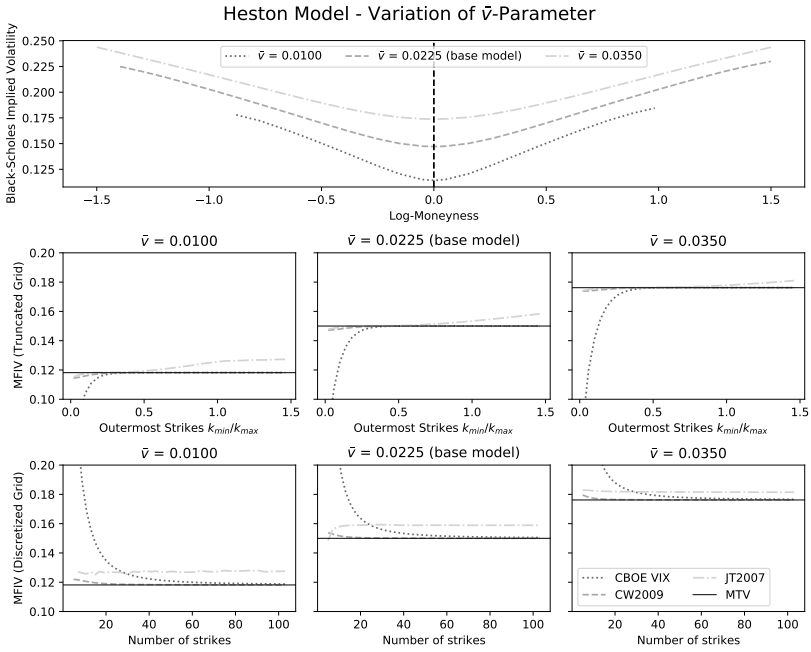

Figure 2.4: Comparison of implementation performance. Option prices have been generated from a Heston model with three parameter sets $\Omega = \{v_0 = 0.0225, \bar{v} = [0.0100, 0.0225, 0.0350], a = 3, \eta = 0.25, r = 0, \rho = 0\}$. Top panel shows implied volatility smile of each parameterization. Middle panels show computed MFIVs for each parameter set after truncation. Bottom panels show computed MFIVs for each parameter set after discretization.

in a highly truncated grid, but converge to MTV quickly and reliably without overshooting. The implementation by Carr and Wu [2009] reduces this effect for a very regular surface. The implementation by Jiang and Tian [2007] takes curvature of the smile at the truncation point into account, therefore the extrapolation overestimates the true wing IV less than on the surfaces presented in figure 2.3.

Figure 2.5 illustrates the error dynamics when varying variance rebound speed κ. With a lower chosen value for κ, the wings of the smile steepen, while the smile remains mostly symmetric in log-moneyness. The patterns observed previously apply here. The methods by CBOE [2018] and Carr

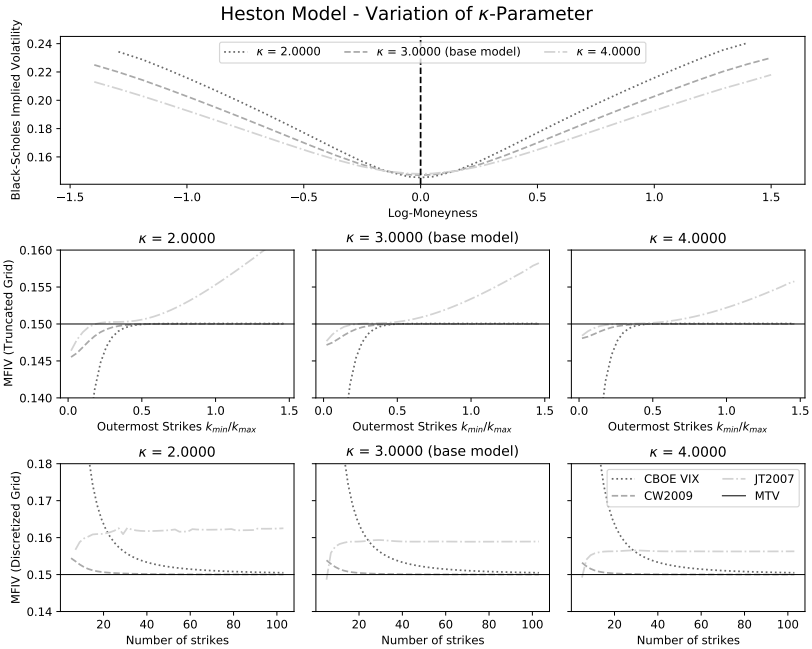

Figure 2.5: Comparison of implementation performance. Option prices have been generated from a Heston model with three parameter sets $\Omega = \{v_0 = 0.0225, \bar{v} = 0.0225, a = [2, 3, 4], \eta = 0.25, r = 0, \rho = 0\}$. Top panel shows implied volatility smile of each parameterization. Middle panels show computed MFIVs for each parameter set after truncation. Bottom panels show computed MFIVs for each parameter set after discretization.

and Wu [2009] exhibit initially large, but reliably converging results. The implementation by Jiang and Tian [2007] overcompensates for truncation, worsening with increased steepness in the wings.

2.6.2 Error Dynamics under Jumps in the Price Process

To illustrate the error dynamics when observing an asymmetric smile, we simulate option prices using a Merton model as described in section 2.5.2.

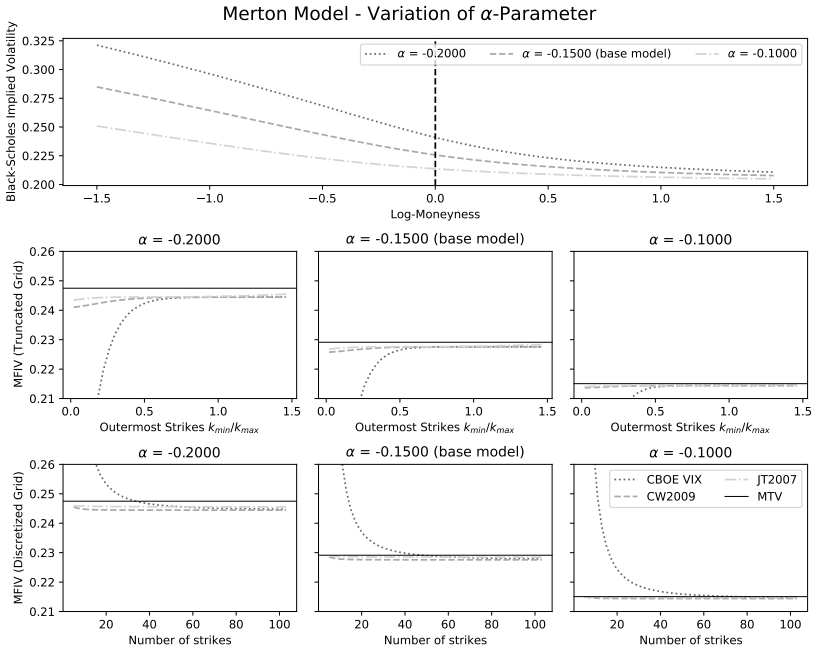

Figure 2.6: Comparison of implementation performance. Option prices have been generated from a Merton model with three parameter sets $\Omega = \{\sigma = 0.2, \alpha = [-0.20, -0.15, -0.10], \delta = 0.05, \lambda = 0.5\}$. Top panel shows implied volatility smile of each parameterization. Middle panels show computed MFIVs for each parameter set after truncation. Bottom panels show computed MFIVs for each parameter set after discretization.

The initial parameterization is $\Omega = \{\sigma = 0.2, \alpha = -0.15, \delta = 0.05, \lambda = 0.5\}$, and individual parameters are modified to change the shape of the volatility smile.

Figure 2.6 shows the error dynamics when changing the mean jump size α. For the chosen parameterization, the mean jump size is negative, and the jump process can be interpreted as sudden market corrections. With larger negative jumps sizes, the option prices in the left wing of the smile are larger, therefore the left wing is steeper. The flatness of the right wing is a consequence of the log-normality of the remaining process. Under these return dynamics, the increased premia in the left wing lead to increased variance

sensitivity. This, in combination with the $\frac{1}{K^2}$-weighting scheme, leads to the general underestimation of MTV by equation 3.2, which is distinct from errors due to discretization and truncation.

While the left wing exhibits positive growth, the right wing is flat. Therefore all three implementations underestimate MTV under artificial truncation. As the smile is less severely truncated, the underestimation diminishes. The implementation by CBOE [2018] is affected strongest, and exhibits the largest underestimation error. The implementation by Carr and Wu [2009] compensates for most of the truncated smile. The log-linear extrapolation scheme suggested by Jiang and Tian [2007] depends on the curvature of the left wing at the respective cutoff point, and overcompensates slightly. Under discretization, both the implementations by Carr and Wu [2009] and Jiang and Tian [2007] compensate for most of the discretized surface. The constant level BSIV interpolation by CBOE [2018] creates a large overcompensation.

2.6.3 Error Dynamics under Stochastic Volatility and Jumps

While the analysis of the error dynamics under Heston- and Merton-style underlyings shed light on the effect of stochastic variance and jump risk in isolation, the analysis of the SVJ model permits insight into the combinatorial properties of the error behaviour. As the two sources of randomness are per equation 2.16 not correlated, we isolate the cumulative effects.

The chosen base parameterization parallels the previous calibrations. The volatility smile consequently exhibits properties of both previously discussed base calibrations. It is upward sloping with flattening tails, but the left tail is steeper and more convex. The global minimum is approximately at $k = 0.2$. As before, the asymmetry in the smile creates a structural discrepancy between MTV and MFIV.

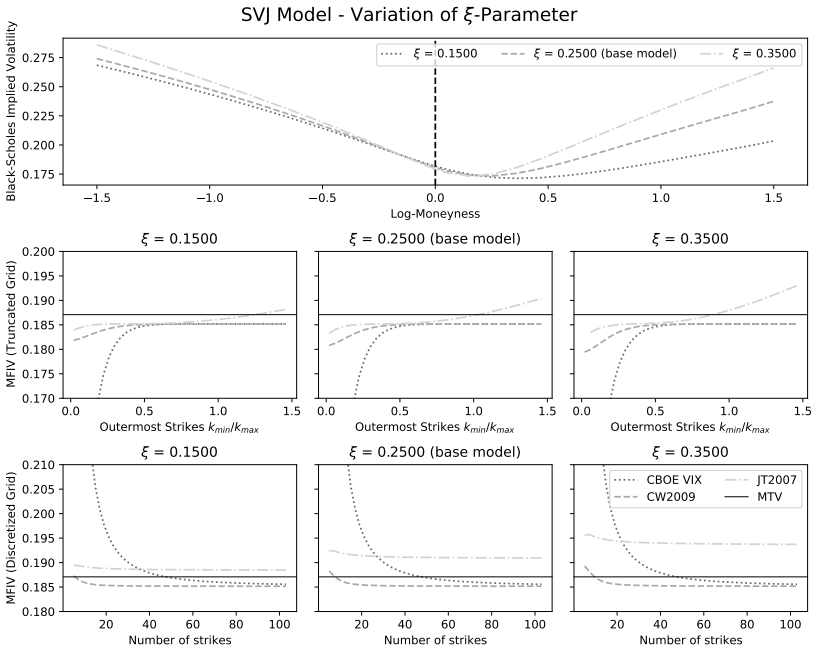

Figure 2.7: Comparison of implementation performance. Option prices have been generated from an SVJ model with three parameter sets $\Omega = \{v_0 = 0.0225, \bar{v} = 0.0225, \kappa = 3, \eta = [0.15, 0.25, 0.35], \rho = 0, \sigma = 0.2, \alpha = -0.15, \delta = 0.05, \lambda = 0.5\}$. Top panel shows implied volatility smile of each parameterization. Middle panels show computed MFIVs for each parameter set after truncation. Bottom panels show computed MFIVs for each parameter set after discretization.

Since the average jump size is negative, the jump component dominates the left wing. Modifying volatility-of-variance parameter ξ of the stochastic variance component has therefore a stronger effect on the right wing than the left wing. Results are provided in figure 2.7. Under truncation, the implementation by CBOE [2018] shows consistent underestimation of MTV. The constant extrapolation approach of Carr and Wu [2009] compensates truncation partially, but requires approximately the same grid width as the implementation by CBOE [2018]. The implementation by Jiang and Tian [2007] compensates best on highly truncated surfaces, but overestimates MTV on wider surfaces. The error size increases with steeper tails. On a discretized smile the CBOE VIX Index implementation overestimates MTV most strongly.

Carr and Wu [2009]'s interpolation leads to small, consistent overestimation and converges quickly. The implementation by Jiang and Tian [2007] compensates discretization, but suffers from consistent overestimation due to extrapolation, consistent with the results from truncated surfaces.

2.7 Summary of Findings

Across all simulations, a common pattern emerges. The implementation by CBOE [2018], which does not employ procedures to adjust for truncated tails, severely underprices forward variance compared to the MTV. The error grows with less option prices observable. The constant extrapolation of Carr and Wu [2009] does reduce the error significantly, but still underprices when very few option prices are observable. Both methods converge reliably to MTV once a sufficiently wide strike domain can be observed. The linear constant slope extrapolation by Jiang and Tian [2007] tends to overcompensate the tails, leading to overpricing of MTV with wider smiles.

Our key result is that the VIX index is very error prone in real-world conditions, and that these errors should be accounted for when using the VIX index as a proxy for market risk aversion. In markets with short option chains, it might be beneficial to use parametric models to estimate MFIV. The presence of jump risk leads to highly asymmetric volatility smiles, increases the magnitude of error, and makes its behaviour less predictable.

Discretization generally leads to larger error magnitudes because it occurs, by definition, closer to ATM than the truncation. The implementation of artificial discretization also affects the resulting error dynamics. For example, Jiang and Tian [2005] choose a different approach to artificially discretize the volatility surface, which, under their proposed interpolation method, leads to fluctuating error signs due to the ATM strike being either under- or overestimated. Our discretization scheme forces the existence of an ATM strike

to account for this.

As the log-linear extrapolation of Jiang and Tian [2007] tends to overestimate BSIV in the tails due to curvature, the truncation error at $k = 2$ also affects the discretization error for all discretization simulations.

2.8 Conclusion

In this chapter we show that the accuracy of MFIV calculations in practice is highly dependent on the choice of interpolation and extrapolation methods. Our findings reveal that the most common features of stochastic models for equities, stochastic variance and jumps, lead to distinct error dynamics for each method. The choice of model should take the length and resolution of the option chain as well as the features of the underlyings' process into consideration.

None of the implementations discussed attempt to generate arbitrage-free volatility smiles. While this tends to average out for interpolation because the surrounding strikes of each interval provide reasonable boundaries, extrapolation without regard for no-arbitrage constraints causes large errors, even on very wide smiles.

Our results are, however, only valid in market configurations that can be represented with our chosen stochastic processes. The reason for choosing these specific models is that they exhibit both stochastic variance as well as jumps, and allow us to compare both in the SVJ model. An interesting question for further research lies in error dynamics under market configurations with large outlier events.

For highly liquid indices with very active and dense option chains and exhibiting smooth volatility surface, a simple approach, such as suggested by CBOE [2018], seems appropriate. For individual stocks with relatively illiq-

uid option markets, the computation of MFIV might not be as model-free as intended, and interpolation and extrapolation introduce model dependency though the back door. These issues should receive special consideration when applying a relatively simple implementation to individual stocks as underlying, such as suggested by CBOE [2011].

Furthermore, this analysis focusses on maturities where the reversion to the mean is dominant. For tactical risk management, it is be of interest to examine the convergence in the short term as well. In that case, one would have to analytically derive the model-true short term variance. Further analysis should also investigate the convergence in a regime-switching world.

Chapter 3

Asymptotic Extrapolation of Model-free Implied Variance

This chapter was published as
Stahl, P.: Asymptotic extrapolation of model-free implied variance: exploring structural underestimation in the VIX Index.
Review of Derivatives Research 25, 315–339 (2022)
https://doi.org/10.1007/s11147-022-09190-2

We show that the VIX Index structurally underestimates model-free implied volatility because its implementation omits extrapolation of the volatility smile in the tails. Hedging strategies that reduce their risk exposure in response to elevated VIX levels are therefore systematically underhedged. We use the asymptotic behavior of the volatility surface to construct a correction term that is model-independent and only requires option prices at the two outermost strikes. We show how to apply this correction to the VIX Index ex-post as well as how to modify its implementation accordingly. Furthermore, we show that the degree of underestimation varies over time. For the S&P

500 Index and the DJIA Index the error is larger in periods of sustained low volatility. This cannot be observed for the Volatility-of-VIX Index. By incorporating strike grid width into the calculation, the suggested correction reduces uncertainty of the volatility index, which may benefit VIX-based hedging stratgies.

3.1 Introduction

The CBOE VIX Index is the one of the most commonly used indicators for investor fear. It implements the static replication approach to variance swap pricing by Carr and Madan [1998], who show how to replicate the future realized variance of an assets' return process between today and a future date T using the prices of options expiring at T. Britten-Jones and Neuberger [2000] coin the term *model-free implied variance (MFIV)*, as the replication strategy does not depend on the assumption of a stochastic model for the underlying. The replication portfolio requires a continuous set of strikes from 0 to infinity. Since in practice option strikes are discrete and only actively traded in a relatively narrow range around the current price of the underlying, systematic errors are introduced.

Specifically, the methodology behind the VIX Index as outlined in CBOE [2018] truncates the tails completely, and makes no effort to compensate for this. This leads to structural underestimation of MFIV in the VIX Index implementation. Option trading activity varies over time with changing market regimes. When interest in far-from-the-money options wanes during calm market periods, the observable area of the volatility smile is truncated further than in more volatile periods. As the width of the observable option strike grid changes, the underestimation varies over time. Many hedging strategies reduce their risk exposure when the VIX Index is elevated. The structural underestimation therefore implies that these strategies underestimate market risk, and are systematically overinvested or underhedged. The

goal of this chapter is to examine this truncation error and to understand its dynamics over time. The truncation error term can compensate for the time-varying structural underestimation in the VIX Index by systematically accounting for grid width, and make values from different market regimes or underlyings comparable.

To compensate for truncation, we apply a result by Lee [2004] on the asymptotic behaviour of the volatility smile. By deconstructing the main integral of Britten-Jones and Neuberger [2000], which is based on Carr and Madan [1998], we isolate the tails of the volatility smile and compute them using asymptotic option prices. This yields a very simple truncation error correction term that is completely independent of model assumptions. The only required observations are the option prices at the outermost strike to parameterize each tail. The calculation is demonstrated using the reference data by CBOE [2018].

We calculate the truncation error for volatility indices of both the S&P 500 Index and the Dow Jones Industrial Average Index, and show how the truncation error varies depending on market conditions. As a counterexample, we show that the Volatility-of-VIX Index does not exhibit this pattern.

Our findings have implications for the analysis of the volatility premium as well as VIX Index futures. The volatility premium refers to the positive expected return that sellers of options capture. In the literature, it is commonly defined as the difference between the volatility of the underlying price and the VIX Index. For example, Eraker [2021] presents a general equilibrium model based on long-run risk that is able to capture the volatility premium as well as the negative correlation between prices of the underlying. Gruber et al. [2020] construct a state-based volatility process to model the term structure of the volatility premium in periods of low and high volatility separately. The structural underestimation of the VIX Index might explain a small portion of this premium. Bardgett et al. [2018] examine the information content of options on VIX Index futures and document a that it varies over time.

Cheng [2019] studies the volatility premium embedded in VIX futures and finds a lower than expected response to rising market risk. Bakshi et al. [2021] show that VIX futures are in contango when jumps in volatility exist, and are correlated with jumps in the price of the underlying. This implies that the volatility premium compensates the option seller for the risk of large jumps. Eraker and Yang [2020] construct a sophisticated consumption-based equilibrium framework to integrate the pricing VIX and SPX options as well as equity and variance premium. The structural underestimation in the VIX Index is likely priced into VIX futures, and the described patterns might distort the VIX Index futures term structure.

To begin our analysis, we review the related literature. Next, we examine the construction of the VIX Index in detail to identify its structural bias. We then investigate the asymptotic behaviour of the volatility smile and show how to correct for truncation. Finally, historical option prices are used to calculate the truncation error over time, revealing its dependence on market volatility.

3.1.1 Notation

Markets are assumed to be free of arbitrage opportunities and complete. This implies the existence of an *equivalent martingale measure (EMM)* which is uniquely characterised by the risk neutral density $\Phi^{\mathbb{Q}}$ of the stochastic process ruling the underlying. The fair value of a derivative is given by the expected value of its payoff under $\mathbb{Q}$.

Contracts live from time 0 to time T, which is specified in years. The current time is $t \in [0, T]$. The time to maturity is $\tau = T - t$. For discrete returns, actual trading days per year are used for annualization; for continuous cases, 365.25 calendar days and 252 trading days are assumed.

The price of an underlying at t is S_t, its forward price at time t is $F_t = S_t e^{r\tau}$.

Option strikes in absolute (dollar) terms are denoted K, while the strike in terms of log-moneyness against the forward price is denoted $k = \log\left(\frac{K}{F_0}\right)$. Capitalized option prices $C(K,\tau)$ and $P(K,\tau)$ are quoted in dollar value, their counterparts in log-moneyness terms, such that $C(K,\tau) = F_0 c(k,\tau)$ and $P(K,\tau) = F_0 p(k,\tau)$. Unless explicitly specified, prices refer to the average of the bid and ask quotes. $Q(K,\tau)$ and $q(k,\tau)$ denote the price of the out-of-the-money option in dollar and log-moneyness, respectively.

The term *implied volatility (IV)* refers to the *Black-Scholes implied volatility (BSIV)* σ_{BSIV} that solves

$$C_{observed}(K,\tau) = C_{Black\text{-}Scholes}(K,\tau,\sigma_{BSIV}) \tag{3.1}$$

uniquely, where $C_{observed}(K,\tau)$ is an observed call price at strike K and $C_{Black\text{-}Scholes}(K,\tau,\sigma_{BSIV})$ is the Black-Scholes price at strike K and volatility σ_{BSIV}. The term *volatility smile* describes the set of BSIVs that is produced by all options on an underlying asset with the same expiration date. The term *volatility surface* refers to the collection of volatility smiles of all available expiration dates at a certain point in time.

Unless explicitly specified, all options are of European style, and dividends and interest rate r are assumed to be zero and omitted for clarity.

3.1.2 Literature Review

Some fundamental results in the literature provide insight into the relationship between observed option prices and forward variance of the underlying. Breeden and Litzenberger [1978] show that the risk neutral price probability distribution of an underlying at expiration is uniquely determined by the complete set of options without assumption of a parametric model for the underlying. Neuberger [1994] introduces the concept of futures contracts that pay the natural logarithm of an underlying future at expiration and shows

that the payoff of a delta-hedged log-contract depends purely on volatility of the underlying, provided its variance is constant over time. Independently, Dupire [1994] shows that a unique risk-neutral density $\Phi^{\mathbb{Q}}$ can be recovered from market prices, if all prices are compatible with no arbitrage conditions and the underlying is governed by a diffusion process.

Carr and Madan [1998] show that a volatility swap whose payoff is the realized volatility of the underlying can be fairly priced by statically replicating log-contracts using the result of Breeden and Litzenberger [1978] and delta-hedging them as suggested by Neuberger [1994]. Specifically, they show that a $\frac{1}{K^2}$-weighted portfolio of out-of-the-money options has virtually constant sensitivity to changes in variance, and construct a portfolio that perfectly replicates forward variance of a continuous underlying[1]. Furthermore, it is only based on the assumption of risk-neutrality, and does not rely on a specific stochastic model for the underlying process, except that it follows a diffusion. Britten-Jones and Neuberger [2000] extend their analysis and show that this result holds in the case of stochastic volatility in the underlying as well. Demeterfi et al. [1999] explore the effect of skewness on the price of variance and volatility swaps for continuously moving underlyings. Based on Carr and Madan [1998], Britten-Jones and Neuberger [2000] define MFIV as

$$\mathbb{E}_0^{\mathbb{Q}}\left[\int_0^T \left(\frac{dF_t}{F_t}\right)^2\right] = 2\int_0^\infty \frac{C(K,\tau) - max(F_0 - K, 0)}{K^2} dK, \tag{3.2}$$

with K being a strike from the set of all observable strikes, and, since the expectation is taken under $\mathbb{Q}$, $C(K,\tau) = \mathbb{E}_t^{\mathbb{Q}}[max(F_t - K, 0)]$.

Jiang and Tian [2005] show that equation 3.2 holds for price processes with jumps. They find evidence that the MFIV is a more accurate predictor of future realized variance than historical variance or Black-Scholes ATM implied variance.

[1] For detailed derivation, see Demeterfi et al. [1999] or Derman and Miller [2016], Ch. 4.

Even though the result by Carr and Madan [1998] has been thoroughly investigated, the discrete and truncated nature of option markets still presents challenges in practical applications. The VIX Index was originally introduced by the *Chicago Board Options Exchange (CBOE)* in 1993 to measure market-expected 30-day volatility of the S&P 100 Index. In its original formulation as suggested by Fleming et al. [1995] it was defined as at-the-money Black-Scholes implied volatility, and computed by solving equation 3.1 with the current 30-day at-the-money option. In 2003, the VIX Index was updated to reflect the advances in option pricing theory laid out above, in conjunction with a change of the underlying to the S&P 500 Index. CBOE [2018] provide details on the updated calculation and section 3.2 discusses its shortcomings. CBOE [2011] introduced a single-stock variant of the VIX Index for a handful of stocks whose options are very liquid. In 2012, CBOE [2012] introduced the Volatility-of-VIX Index, which applies the methodology of the VIX Index to options on the VIX Index itself. Jiang and Tian [2005] examine the practicality of 3.2 and identify two distinct sources of implementation error: discretization due to the fact that observed strikes of options are naturally discrete, but equation 3.2 is based on a continuous integral in K; and truncation because equation 3.2 requires integration along $K \in [0, \infty)$. They state theoretical upper bounds for the truncation and discretization errors. Jiang and Tian [2007] examine the implementation of equation 3.2 by CBOE [2018] and find it to be systematically flawed. They show that, depending on the volatility environment, the magnitude of the implementation error is predictable. This chapter partially extends their analysis. They suggest a numerical scheme to overcome the issues using cubic splines for interpolation, and linear extrapolation in log-moneyness space k beyond the outermost strikes. Benaim et al. [2009] investigate the concept of model-based interpolation and extrapolation, where a stochastic process is calibrated to fit option prices, which is used to generate a synthetic option price surface that overcomes the issues of truncation and discretization. They find that this approach introduces systematic errors in the tails. They suggest supplementing a model-based interpolation with numerical extrapolation. Broadie and Jain [2008b] analyze the effect of discretization on

the pricing of variance swaps using a variety of stochastic processes for the underlying and find that errors due to discretization are usually small, but the effect of jumps – which manifest in the tails of the smile – can be large. Carr and Wu [2009] and Jiang and Tian [2005] extrapolate with constant BSIV beyond the outermost strikes. While simple, this approach has several drawbacks. First, it satisfies the conditions set forth by Benaim and Friz [2009] if and only if we limit the underlying process to be log-normally distributed (the Black-Scholes case). For this specific case, the approach outlined in this chapter can be adapted using the distribution-specific dynamics of Benaim and Friz [2009] to achieve an equivalent result. Second, it introduces dependency on the observed cutoff point: a cutoff point at a larger $|k|$ has, under the described growth dynamics, necessarily a equal or larger BSIV, such that constant extrapolation leads to similar, albeit lower, structural underestimation as the implementation by CBOE [2018]. Jiang and Tian [2007] furthermore criticise the introduction of kinks into the volatility smile, which violate no-arbitrage conditions. They interpolate with natural cubic splines and extrapolate BSIV linearly in log-moneyness, with the slope of the extrapolation function matching the first derivative of the spline at the outermost strike. Except for a Black-Scholes world where a constant extrapolation would be exact, any linear growth violates the no-arbitrage bounds set forth by Lee [2004][2]. Carr and Wu [2009] utilize MFIV to estimate variance risk premia for indices and single stocks. They also provide a detailed interpolation and extrapolation procedure, which differs from Jiang and Tian [2007] in that they use linear interpolation between strikes and constant extrapolation beyond the outermost strikes.

The relationship of the distribution of the underlying and the shape of the volatility surface permits analysis of the asymptotic behaviour of BSIV in the tails. Hodges [1996] establishes that the no-arbitrage bounds set forth by Merton [1973] can be expressed by quoting option prices in terms of

[2]Any positive linear growth violates Lee [2004]'s upper bound, and any negative linear growth is excluded by the lower bound of constant growth in the Black-Scholes case, where all moments of the underlying distribution are finite. For details, see Gatheral [2006, p. 99]

their BSIV, where a positive BSIV prevents arbitrage of between options and the underlying plus cash, and having a single BSIV per strike and T enforces Put-Call parity. Furthermore, he provides bounds for the slope of implied volatility in the tails. Lee [2005] describes the static and dynamic characteristics of the volatility smile, and shows how the volatility smile can be interpreted as probabilistic density. Carr and Wu [2016] provide insights into the dynamic evolution of the volatility surface and derive no-arbitrage constraints based on those dynamics. Lee [2004] examines the asymptotic behaviour of BSIV in the strike domain. He finds that under absence of arbitrage the growth of BSIV is bound from above by $\sqrt{\frac{\beta}{T}|k|}$, where $\beta \in [0, 2]$ being specific to either the left or right wing. Benaim and Friz [2009] expand and refine this result and show that the asymptotics of the volatility smile are a non-linear transform of the asymptotics of the underlyings' return distribution. They show that the result of Lee [2004], under some mild technical conditions, precisely determines tail behaviour. Furthermore, they show how to explicitly derive the asymptotic behaviour of the BSIV for a variety of stochastic models. Benaim et al. [2012] examine the relationship between the moment-generating function and the moment formula. They suggest using it to extrapolate the implied volatility surface in the strike domain. In a preceding analysis, Drăgulescu and Yakovenko [2002] come to agreeing results for the special case of distributions with stochastic variance. Gulisashvili [2010] provides asymptotic formulas for call options, as well as error estimates, based on Lee [2004] and Benaim and Friz [2009].

3.2 Structural Shortcomings of the VIX Index

The calculation of the VIX Index is laid out in detail in CBOE [2018]. An overview is provided in appendix A.1.

CBOE [2018] implement equation 3.2 as a weighted Riemann sum based on Demeterfi et al. [1999] using the midpoint rule as

$$MFIV^2 = \frac{2}{T}\Sigma_{i=1}^{N}\frac{\Delta K_i}{K_i^2}e^{rT}Q_{K_i} - \left(\frac{F_t}{K^*} - 1\right)^2, \tag{3.3}$$

with $N \in \mathbb{N}^+$ being the number of observed option prices, K_i being a strike price with $K_i < K_{i+1} \forall i \in \mathbb{N}^+$, K^* being the at-the-money strike price, Q_{K_i} being the observed price of an out-of-the-money option at K_i, $\Delta K_i = \frac{K_{i+1}-K_{i-1}}{2} \forall 1 < i < N$ for all strikes between endpoints, and $\Delta K_1 = K_2 - K_1$ and $\Delta K_N = K_N - K_{N-1}$ for the strikes at the endpoints.

Equation 3.3 implies that the tails of the volatility smile are cut off, leading to a systematic underestimation of forward variance.The truncation error becomes larger when the strike price grid becomes narrower in k. It also grows after large price movements in the underlying, specifically in the time between large price drops in the underlying, and the creation of new options by market makers.

Between strikes, each observation is weighted by half of the difference of the surrounding strikes. This is akin to a constant level interpolation symmetrically around each observation. The errors of this interpolation between strikes tend to be approximately self-cancelling, leading to a small error from interpolation. This does not happen in the region around the strike with the smallest Black-Scholes implied variance, leading to a fluctuating error sign. Some discussion on this can be found in Jiang and Tian [2005][3]. The extend of this depends on the shape of the smile around its minimum. Higher skewness in the underlying implied distribution tend to increase the fluctuation.

[3]Specifically, see figure 2 in Jiang and Tian [2005].

3.3 Asymptotic Extrapolation and Construction of the Truncation Error

The implementation of the VIX Index by CBOE [2018] does not compensate for truncation. By deconstructing equation 3.2, we compensate for the missing tails. We then examine the required strike grid width as well as convergence behaviour of the truncation error.

3.3.1 Construction of the Truncation Error

To compensate for truncation we utilize the asymptotic behaviour of BSIV by computing the respective Black-Scholes option prices and calculating the partial MFIV contribution of each tail.

Extrapolating with constant BSIV of the outermost option price would make the truncation error compensation highly dependent on only the price of these options, without taking the varying strikes into account. We utilize Gulisashvili [2010], who shows that Lee [2004]'s formula can be used to extrapolate option prices in the strike domain.

If we where to assume knowledge of the underlying distribution, we could use the precise tail behaviour based on Benaim and Friz [2009], as suggested by Gulisashvili [2010]. The specific case of a Black-Scholes-compliant underlying would then result in constant extrapolation. While the boundary condition of Lee [2004] does not necessarily apply close to the money, it allows us to find the non-parametric asymptotic truncation error of MFIV, provided $k_{min} \leq 0 \leq k_{max}$ and both k_{min} and k_{max} have sufficient distance from 0. Section 3.3.2 examines this issue in greater detail.

To compute the truncation error compensation, we first deconstruct MFIV into three segments: the observed center segment, and two tails. For the

tails, we derive the Black-Scholes option price as a function of k and a tail-specific parameter β. After calculating β for each tail to fit the outermost option price, we substitute the option price within the tail segments of the deconstructed MFIV. The center segment is left unchanged. The two tails compensate the missing MFIV contribution of equation 3.3.

By rewriting equation 3.2 in log-moneyness terms, we get

$$MFIV = 2 \int_{-\infty}^{+\infty} \frac{c(k, \tau, \sigma_k) - max(0, 1 - e^k)}{e^k} dk. \tag{3.4}$$

Appendix A.2 provides a detailed derivation.

Since the asymptotic BSIV depends exclusively on k and β as $k \longrightarrow \pm\infty$, we can define the extrapolated price of a far-out-of-the-money call option $\tilde{c}(k, \tau, \beta)$ as

$$\tilde{c}(k, \tau, \beta) = \Psi(\tilde{d}_1(k, \tau, \beta)) - e^k \Psi(\tilde{d}_2(k, \tau, \beta)), \tag{3.5}$$

$$\tilde{d}_1(k, \tau, \beta) = -\frac{k}{\sqrt{\beta|k|}} + \tfrac{1}{2}\sqrt{\beta|k|}, \text{ and} \tag{3.6}$$

$$\tilde{d}_2(k, \tau, \beta) = \tilde{d}_1(k, \tau, \beta) - \sqrt{\beta|k|}. \tag{3.7}$$

where Ψ is the Normal CDF [4].

As suggested by Benaim et al. [2012], β is chosen according to the outermost observed option for each tail individually. Since

$$\beta := \limsup_{k \longrightarrow \pm\infty} \frac{\sigma_{BS}^2 T}{|k|}, \tag{3.8}$$

we fix the outermost strike for each tail, and compute the respective $\beta_{left,right}$

[4] See Appendix A.3 for derivation.

as

$$\beta_{left} = \frac{T}{|k_{min}|}\sigma^2_{k_{min}} \text{ and} \tag{3.9}$$

$$\beta_{right} = \frac{T}{|k_{max}|}\sigma^2_{k_{max}}. \tag{3.10}$$

First, we derive the right side truncation error where $k > 0$. The observation farthest to the right is located at k_{max}. Hence, $max(0, 1 - e^k) = 0$. The truncation error to the right side is a function of only k_{max} and β_{right} given by

$$TE_{right}(k_{max}, \beta_{right}) = 2\int_{k_{max}}^{+\infty} \frac{\tilde{c}(k, \tau, \beta_{right})}{e^k} dk. \tag{3.11}$$

The left side truncation error follows in similar fashion, using the asymptotic price of a far-out-of-the-money put option such that

$$\tilde{p}(k, \tau, \beta_{left}) = \tilde{c}(k, \tau, \beta_{left}) - (1 - e^k) \text{ and} \tag{3.12}$$

$$TE_{left}(k_{min}, \beta_{left}) = 2\int_{-\infty}^{k_{min}} \frac{\tilde{p}(k, \tau, \beta_{left})}{e^k} dk. \tag{3.13}$$

The total truncation error is the sum of both sides given by

$$\begin{aligned} TE_{total}(k_{min}, k_{max}, \beta_{left}, \beta_{right}) &= TE_{left}(k_{min}, \beta_{left}) \\ &\quad + TE_{right}(k_{max}, \beta_{right}) \qquad (3.14) \\ &= 2\int_{-\infty}^{k_{min}} \frac{\tilde{p}(k, \tau, \beta_{left})}{e^k} dk \\ &\quad + 2\int_{k_{max}}^{+\infty} \frac{\tilde{c}(k, \tau, \beta_{right})}{e^k} dk. \qquad (3.15) \end{aligned}$$

In comparison to Jiang and Tian [2005], who provide parametric as well as non-parametric upper bounds for the truncation error, this result provides an asymptotic value for the total truncation error. Provided k_{min} and k_{max}

are sufficiently far from 0, this yields a correction term that can be added to MFIV without relying on numerical methods, and instead only relying on asymptotic properties of the volatility surface.

To calculate this truncation error, the uncorrected MFIV value, the observed minimum and maximum option prices, and their strikes are required. Beyond absence of arbitrage, no further parametric assumptions are required.

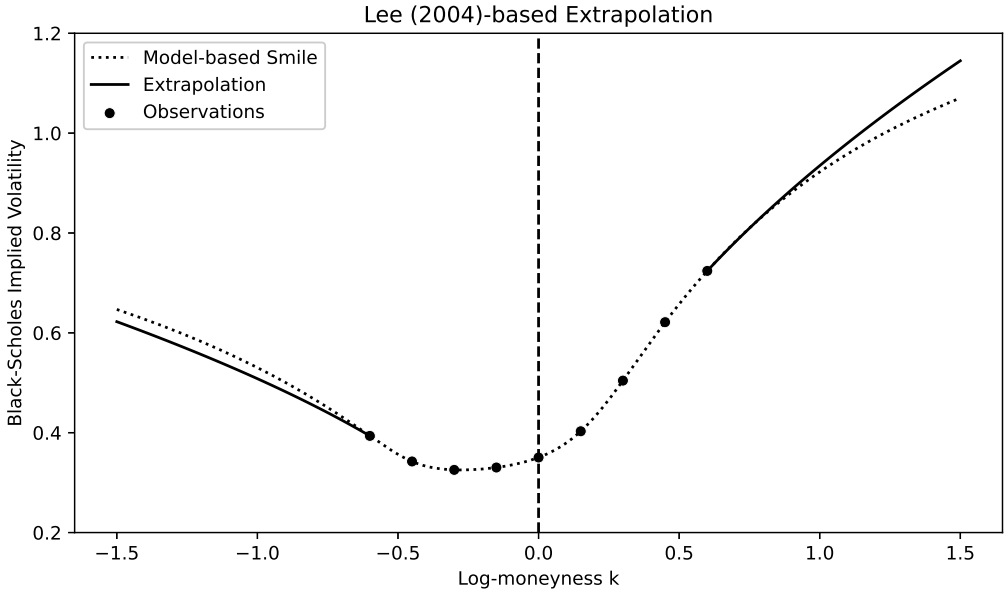

Figure 3.1: Illustration of the extrapolation approach. Simulated volatility smile generated with a Merton-model with $\Omega = \{\sigma = 0.3, \alpha = 0.5, \delta = 0.5, \lambda = 0.1\}$. Left wing cutoff point $k_{min} = -0.6$ with $\beta_{left} = 0.0646$, and right wing cutoff point $k_{max} = 0.6$ with $\beta_{right} = 0.2185$. Note that the extrapolation extends beyond the shown range and the extrapolation is only limited by machine precision.

Figure 3.1 illustrates the extrapolation approach on a simulated volatility smile. The dotted line shows the volatility smile of a Merton-model with the specified parameters. The cutoff points are chosen to be $k_{min} = -0.6$ and $k_{max} = 0.6$, and we calculate β_{left} and β_{right} as described above. The extrapolated wings are shown as a solid line. In the implementation described by CBOE [2018], the extrapolated wings are considered to be 0.

To apply the correction to an observed volatility index VI, square it to convert

to variance, add the error correction from 3.14, and take its square root

$$VI_{corrected} = \sqrt{VI^2_{uncorrected} + TE_{left}(k_{min}, \beta_{left}) + TE_{right}(k_{max}, \beta_{right})}. \tag{3.16}$$

3.3.2 Minimum Requirements for Strike Grid Width

The result of Lee [2004] only holds asymptotically, therefore the suggested truncation error may lead to overestimation when the observable strike grid is too narrow.

To confirm the applicability of the extrapolation, we analyse the behaviour of the tail-parameters β_{left} and β_{right} as $|k| \to \infty$ and show that the observed strike grid tends to be wide enough to admit extrapolation. A reasonable minimum cutoff strike can be found be found by calculating β for every possible cutoff strike of the volatility smile. By definition, β will level off further out-of-the-money, where the extrapolation will be in agreement with the smile. Since the observed volatility smile is truncated, this requires modelling the underlying explicitly. Gulisashvili [2010] provides explicit error estimates for the extrapolation term if the model is known. By imposing a minimum cutoff strike for each wing, this uncertainty of the extrapolation can be reduced.

The VIX Index has a time-to-expiration of 30 days. For short expirations, the jumps tend to be a more important feature than stochastic volatility. Therefore the Merton jump-diffusion model has been chosen as illustrative benchmark[5].

The top panel of figure 3.2 shows the dynamics of β_{Left} and β_{Right} as the cutoff strike is shifted outwards for multiple expirations. With increasing time to maturity, leveling off happens slower and the grid width requirement

[5]Other models and parameterizations are provided in appendix A.4.

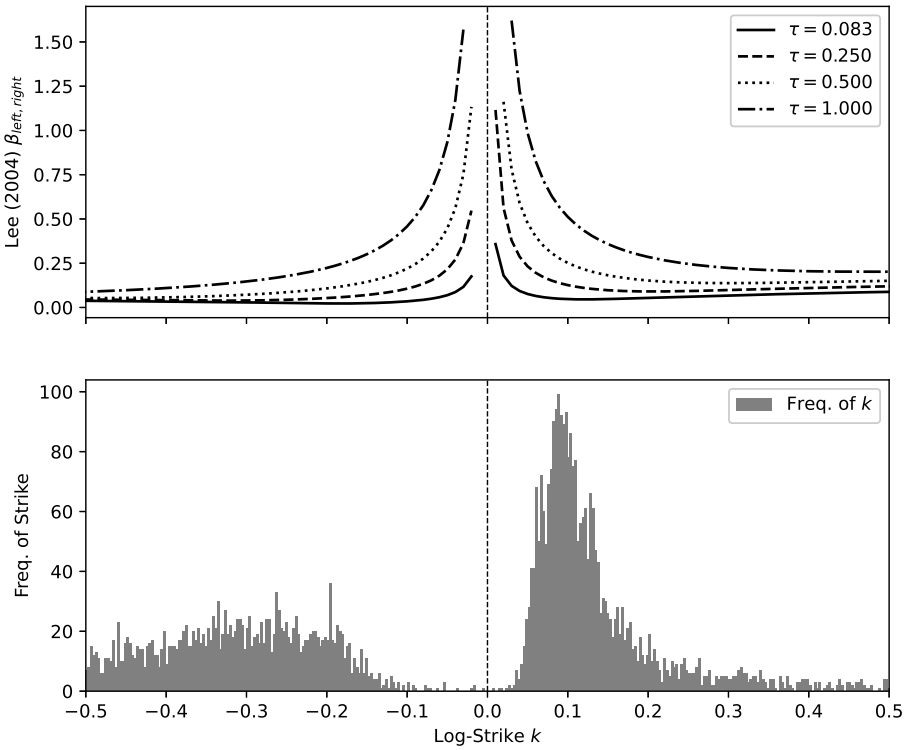

Figure 3.2: Top panel shows the dynamics of Lee [2004]'s tail parameter β as a function of k for various expiration times. Computation based on a Merton jump-diffusion model with parameter set $\Omega_{Merton} = \{\sigma = 0.2, \alpha = -0.15, \delta = 0.05, \lambda = 0.5\}$ (Calibration parameters for the S&P-500 Index from Gatheral [2006, p. 63]). Bottom panel shows the frequency of daily cutoff strikes of the SPX Index with a time-to-maturity between 28 and 32 days (approximately $\tau = 0.083$) between January 1996 and April 2016. Data extends beyond the shown range, but is hidden for visual clarity. Results for different models and parameterizations are provided in appendix A.4

widens. Depending on data availability, the approach by Lee [2004] might not be sufficient. The model-specific approach by Benaim and Friz [2009] may help alleviate some uncertainty of the estimators, however one would be left with model risk.

The bottom panel of figure 3.2 shows a histogram of cutoff strikes for both left and right wing for SPX Index. Based on this data, taking the intended $\tau = 0.083$ into consideration, a minimum cutoff strike of $k = \pm 0.05$ appears to provide a reasonable compromise between data availability and resulting uncertainty for our application. For the SPX Index, this condition is satisfied for 99.8% of data points in the left wing, and for 97.8% of data points in the right wing[6].

3.3.3 Convergence of TE_{Left} and TE_{Right}

Note that the integral in equation 3.2 is well-defined. By the substitution rule, the integral in equation 3.4 is also well-defined. In particular, we have

$$\int_{\tilde{k}}^{\infty} \frac{c(k,\tau,\sigma_k) - max(1-e^k, 0)}{e^k} dk \to 0, \text{ as } \tilde{k} \to \infty \tag{3.17}$$

and the same is true for the integral from $-\infty$ to $-\tilde{k}$.

By definition of $\tilde{c}(k,\tau,\beta)$ and $\tilde{p}(k,\tau,\beta)$ and the asymptotic behaviour of σ_k by Lee [2004], we can approximate TE_{Total} to any given degree by moving the boundary for the approximating proper integral further out, and are limited only by machine precision[7].

[6] The results of the SPX Index analysis for different minimum cutoff strikes are provided in appendix A.5.

[7] Our analysis is implemented in `Python 3.9 x64` and appears to be stable to approximately $k = \pm 30$, which should suffice for any practical application.

3.4 Applying the Truncation Error Term to the VIX Index

To apply the truncation error to the VIX Index we must either modify the VIX Index to remove the extrapolation described above or calculate the actual cutoff points for the truncation error. Using the sample data provided by CBOE [2018] we show how to practically apply the truncation error.

3.4.1 Preliminary Modifications

By construction, the VIX Index implements a very short extrapolation through the definition of ΔK_1 and ΔK_N in equation 3.3. To deal with this modification, we can either shift the cutoff-strikes outward, or modify the underlying price weights at the endpoints.

Ex-Post Approach

When applying our correction term to the VIX Index (and similar indices with this protruding extrapolation), the outermost strikes need to be shifted outwards as

$$K^*_{min} = K_1 - \frac{K_2 - K_1}{2} \text{ and} \tag{3.18}$$

$$K^*_{max} = K_N + \frac{K_N - K_{N-1}}{2}. \tag{3.19}$$

Direct Approach

A slight modification to equation 3.3 can make the outward shift of the cutoff point superfluous. It is then straightforward to include the extrapolation. By

redefining ΔK_{min} and ΔK_{max}, we can avoid the shift described above.

In the original form, the observation at K_i is weighted with $\Delta K_i = \frac{K_{i+1} - K_{i-1}}{2}$, i.e. half of the distance between the surrounding strikes. This implies a constant option price for an interval with length ΔK_i, centered around K_i. At the outermost strikes, $\Delta K_{min,max}$ is defined as the distance to the adjacent strike, e.g. $\Delta K_{min} = K_{min+1} - K_{min}$ ([CBOE, 2018, p.8]).

To simplify the extrapolation, we redefine $\Delta K_{min,max}$ to represent half of the distance to the adjacent strike as

$$\Delta K^*_{min} = \frac{K_{min+1} - K_{min}}{2} \text{ and} \tag{3.20}$$

$$\Delta K^*_{max} = \frac{K_{max} - K_{max-1}}{2}. \tag{3.21}$$

This modification alleviates the implicit extrapolation, but requires computing the center part of the MFIV implementation.

3.4.2 Step-by-Step Example

We will illustrate the approach using the dateset provided by CBOE [2018, Appendix 1 and 2]. The original document provides two option chains, one 25 days from expiration ("near term"), a second chain 32 days from expiration ("next term"). We will show the calculations based on the near term set, and provide results the next term.

We calculate the individual strike contributions as $\frac{\Delta K_i}{K_i^2} e^{rT} q_i$ based on the modified ΔK values as shown in table 3.1. The contributions of the two outermost strikes are halved. We compute the implied variance as

$$\sigma^2 = \frac{2}{T} \left(\sum_i \frac{\Delta K_i}{K_i^2} e^{rT} q_i - \left(\frac{F}{K_0} - 1 \right)^2 \right). \tag{3.22}$$

			Original		Modified	
K	Put/Call	Mid-quote	ΔK	Contribution	ΔK^*	Contribution
1370	Put	0.2	5.0	0.0000005328	2.5	0.0000002664
1375	Put	0.125	5.0	0.0000003306	5.0	0.0000003306
1380	Put	0.15	5.0	0.0000003938	5.0	0.0000003938
	⋮			⋮		⋮
2095	Call	0.2	5.0	0.0000002278	5.0	0.0000002278
2100	Call	0.1	5.0	0.0000003401	5.0	0.0000003401
2125	Call	0.1	25.0	0.0000005536	12.5	0.0000002768

Table 3.1: Strike contributions to CBOE-VIX computation with standard and adjusted definition of $\Delta K_{min,max}$. Affected contributions are highlighted. Data is taken from [CBOE, 2018, Near term data only]

Next, we compute single-sided truncation error corrections using equations 3.13 and 3.11 and add both to the computed variance. Finally, we compute the time-weighed average of the near-term and the far-term variance, and annualise to find the final extrapolated index value. Table 3.2 provides intermediary results.

CBOE [2018] reports a VIX level of 13.69. After the adjustment and compensation for truncation errors on both tails, we find a level of 14.07. In this specific example, truncation leads to an underestimation of variance of approximately 0.38 percentage points.

This adjustment simplifies the extrapolation significantly. The drawback is that it cannot be applied to a precomputed or observed volatility index value, and is therefore mostly useful where MFIV is to be computed from scratch.

3.5 Historical Analysis

The truncation error is larger when the observed option strike grid is narrower. Following demand of market participants, market makers create new contracts. This changes the width of the strike grid over time. In this sec-

	Near-Term	Next-Term
$\sigma^2_{Reference}$	0.018463	0.018821
$\sigma_{Reference}$	0.135878	0.137190
$VIX_{Reference}$	0.136858	
$\sigma^2_{Adjusted}$	0.018479	0.018793
$\sigma_{Adjusted}$	0.135938	0.137087
$VIX_{Adjusted}$	0.136797	
β_{Left}	0.085886	0.081216
β_{Right}	0.059768	0.062062
TE_{left}	0.000542	0.000273
TE_{right}	0.000867	0.000695
$\sigma^2_{incl.TE}$	0.019888	0.019761
$\sigma_{incl.TE}$	0.141025	0.145074
$VIX_{incl.TE}$	0.140688	

Table 3.2: Step-by-step results of ΔK-adjustment and truncation error correction in reference to CBOE [2018].

tion, we analyze how the availability of observable option prices affects the truncation error over time. In periods of high market volatility, the strike grid of observable option prices tends to be wide, which implies a low total truncation error. As volatility levels quiet down after periods of steady growth, the strike grid shrinks, and the truncation error grows larger. In effect, the VIX Index is typically precise in turbulent market phases, but underestimates MFIV when it is low. This effect is shown for the S&P 500 Index (SPX Index) and the associated VIX Index, as well as the Dow Jones Industrial Average (DJIA Index) with its respective volatility index. The Volatility-of-VIX-Index (VVIX Index) does not exhibit this pattern.

3.5.1 Dataset and Computation

The dataset consists of options traded on the CBOE Options Exchange between January 1st, 1996 and April 29th, 2016, spanning 5118 trading days. It contains daily bid and ask prices for 233186 option contract on the SPX Index, the DJIA Index, and the VVIX Index. The options have been selected

to have a remaining lifetime between 23 and 37 days. The same liquidity requirements as laid out in CBOE [2018] are applied. Contracts that have not been traded on a given day (and thus have a "stale price") are excluded for that day. For each day and maturity, once two contracts with neighbouring strikes and stale prices are encountered, all contracts further away from the money are discarded as well. These constraints leave around 2.73 million observations, 1.57 million of which for contracts on the SPX Index. The option price data is matched to price data of the underlying price and to linearly interpolated U.S. treasury rates.

Since the choice of each wings' β depends on the respective outside strike, we require a minimum log-moneyness of ± 0.05. If the strike grid is narrower than this, the truncation error might grow unreasonably large. Based on the analysis in section 3.3.2 and the intended time-to-expiration of 30 days, ± 0.05 appears to provide sufficient space for the β to level out.

For each day with sufficient option price data, the volatility index is calculated in two ways. First, we follow the reference implementation by CBOE [2018]. Second, the truncation error is calculated and added to the strike-adjusted reference implementation. β_{Left} and β_{Right} are computed for the left and right wing based on the price with the lowest (highest) available strike after applying liquidity requirements. This is done for each trading day with available and admissible data.

3.5.2 Analysis

As strike price grid width changes over time, the total truncation error fluctuates. From day to day, fluctuations are due to the discrete nature of strikes and the fact that stale prices are not considered in the calculation. Using smoothed data reveals the structural effect on the volatility index. The truncation error grows in calm market phases, and contracts during correction phases. As market volatility spikes during a correction, high implied

volatility levels and option delta create short-term spikes in the truncation error, which level off quickly. The daily log changes of the corrected volatility indices exhibit lower standard deviation and higher kurtosis than their uncorrected counterparts.

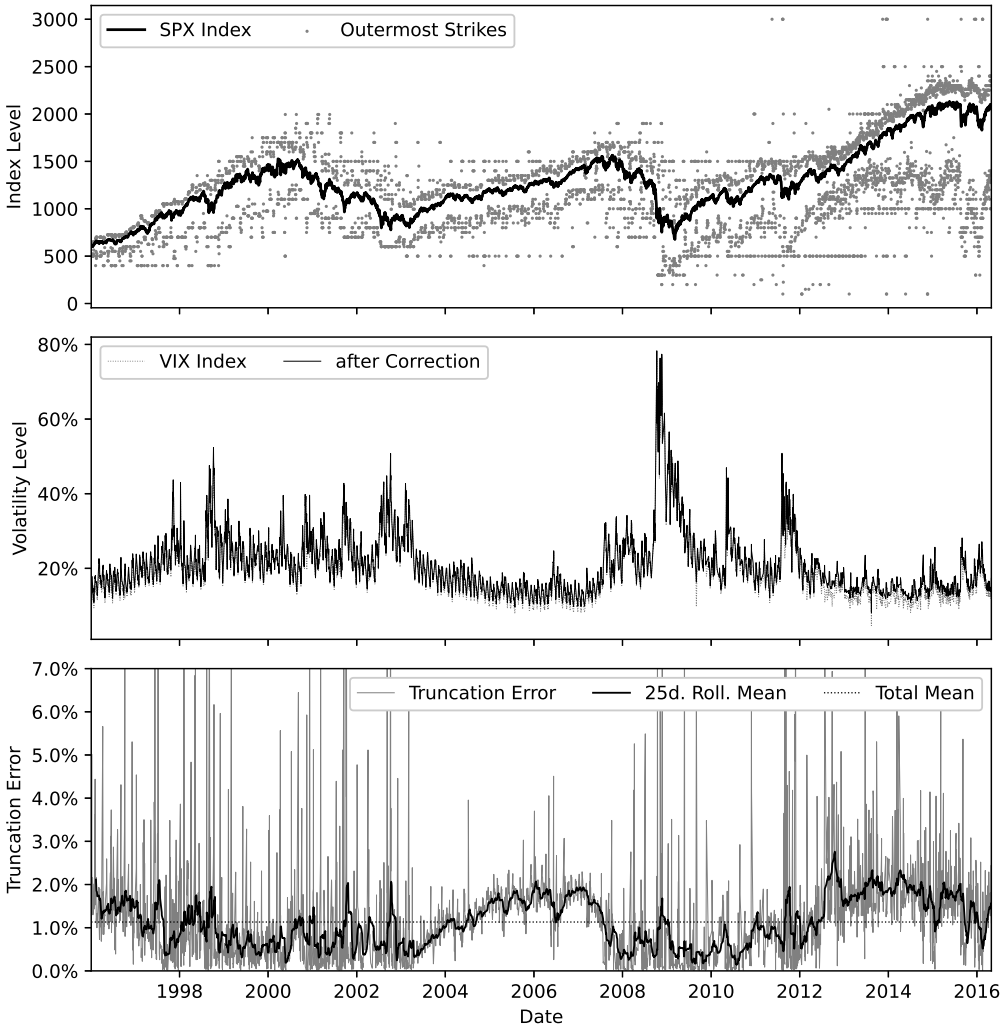

Figure 3.3: Observable option strikes and the effect on the truncation error over time. Top panel shows the S&P 500 Index level and the outermost observable strikes of each respective date. Middle panel shows the VIX Index and the VIX Index after adding the correction. Bottom panel shows the total truncation error and its 25-day rolling mean.

Figure 3.3 shows the behaviour of the truncation error over time. The top

panel illustrates the configuration of the option market in relationship to the underlying SPX Index over time. Each dot represents the outermost option strike observable at each point in time, with the options being subject to the liquidity constraints of CBOE [2018]. The middle panel compares the VIX Index implementation as described in section 3.2 with the amended implementation as described in section 3.4.1. The bottom panel provides the relative error size and its 25-day rolling mean. The overall average truncation error is 1.13%.

In broad strokes, four distinct periods can be identified from figure 3.3. From 1996 to 2004, the SPX Index shows a large correction. The VIX Index is generally above 20%. The truncation error is fluctuating heavily, but its 25-day rolling mean mostly remains below its global average. Between 2004 and 2007 the SPX Index developed positively. The VIX Index mostly remaines below 20%. The truncation error fluctuates less, and its rolling mean rises to 2% and remains on this level. From 2007 to late 2012, the multiple large corrections in the SPX Index take place. The VIX Index spikes during the corrections, and remains above 20%. In this period, the truncation error falls below 1%, where it remains for most of the time. After 2012 until the end of the dataset in 2016, the SPX Index develops positively, and the VIX Index again remains around or below 20%. The truncation error grows to 2% and remains on this relatively high level. In the last few months of the dataset, a slight increase in the VIX Index coincides with a drop in the truncation error.

Overall, the truncation error appears negatively related to the VIX Index. In periods of high volatility, the truncation error fluctuates, but its rolling mean remains below the overall average truncation error. As the VIX Index remains on low levels, the truncation error is consistently elevated.

Figure 3.4 shows the results of the historical calculation for the DJIA Index. Strike price grids, after applying liquidity requirements, for the DJIA Index tend to be narrower than for the SPX Index. The truncation error has an

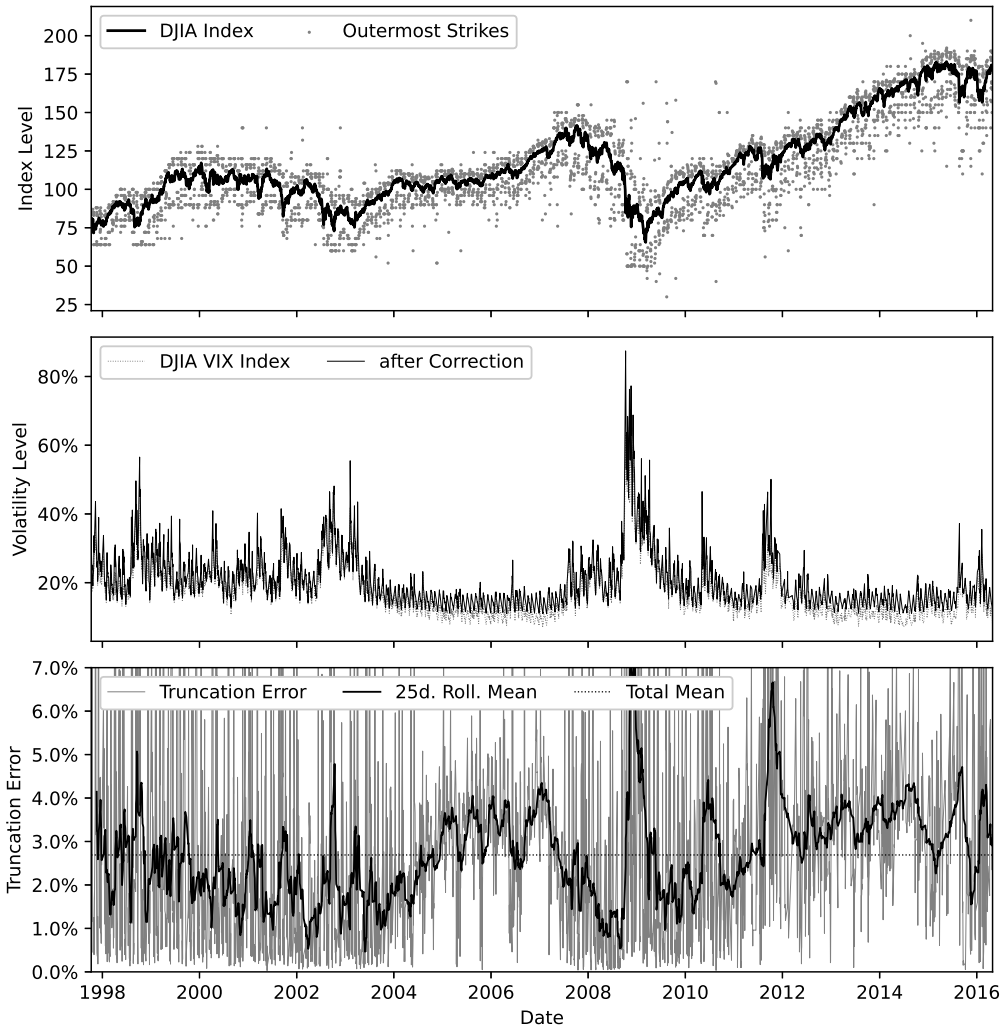

Figure 3.4: Observable option strikes and the effect on the truncation error over time. Top panel shows the DJIA Index level and the outermost observable strikes, subject to liquidity requirements. Middle panel shows the DJIA VIX Index and the DJIA VIX Index after adding the correction. Bottom panel shows the total truncation error and its 25-day rolling mean.

overall average of 2.72%, which is higher than the truncation error of the SPX Index.

The observation period can again be divided into four distinct periods. From the 1996 to 2004, the DJIA Index moves roughly sideways and exhibits ele-

vated levels of volatility including several sharp spikes. The truncation error fluctuates strongly and its rolling mean fluctuates between 1% and 5%, but remains mostly below its overall average. From 2005 to 2007, the DJIA Index exhibits positive performance, and volatility levels remain consistently below 20%. The truncation error in this period is consistently elevated between 2.5% and 4.3%. In the period from 2008 to 2013, coinciding with negative returns, three large volatility spikes can be observed. The truncation error spikes during periods of negative market returns, and falls quickly below its overall average. From 2013 to 2016, DJIA Index returns are positive, and volatility levels are low. The rolling mean of the truncation error remains above its overall average for almost the entire time.

This behaviour is consistent with the analysis of the SPX Index. In periods with low DJIA VIX Index levels, the truncation error tends to be on a higher level. However, immediately after strong market corrections with corresponding volatility spikes, the truncation error spikes as well. This pattern is only very faintly visible in the analysis of the SPX Index.

Figure 3.5 shows the truncation error of the VVIX Index over time, the option-implied forward volatility of the VIX Index. The VVIX Index, detailed in CBOE [2012], applies the same methodology as CBOE [2018] to options on the VIX Index. The VIX Index is governed by a return distribution that is fundamentally different from a stock price index, with the most prominent difference being its mean reversion property in combination with a positive expected jump size. Furthermore, its negative correlation with the SPX Index makes it possible to use options on the VIX Index as hedging instruments for SPX Index-related delta risks. Fernandes et al. [2014] attempt to model the underlying process of the VIX Index directly and confirm both properties.

The dynamics of the VIX Index and the related VVIX Index appear very stable. While there are large spikes in the VIX Index during market corrections in the SPX Index, the VVIX Index exhibits much higher levels than the VIX

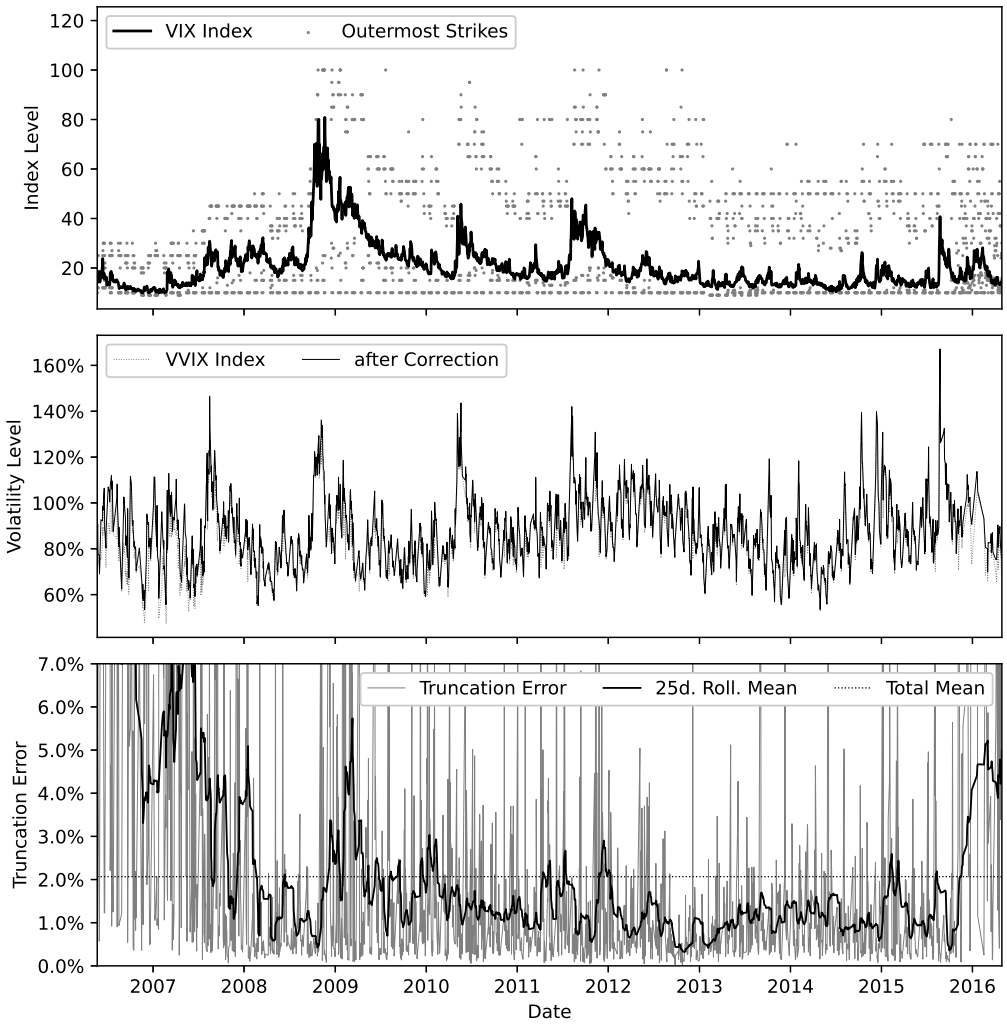

Figure 3.5: Observable option strikes and the effect on the truncation error over time. Top panel shows the VIX Index level and the outermost observable strikes, subject to liquidity requirements. Middle panel shows the VVIX Index and the VVIX Index after adding the correction. Bottom panel shows the total truncation error and its 25-day rolling mean.

Index, and fluctuates much more. The fluctuations are smaller in magnitude than those of the VIX Index, thus its range tends to be more localized over time. The individual trajectories appear more erratic than in the VIX Index, where spikes can often be linked with SPX Index corrections. Spikes in the VVIX Index occur but are less extensive.

Considering the truncation error, the available timeseries can be split into two periods. From 2006 to 2010, the truncation error starts on a very high level, but quickly shrinks to below the overall average level of 2.04%. The market correction in 2009 leads to a spike in the VIX Index, and to a spike in the truncation error levels. In the period after 2010, the truncation error remains mostly below its overall average, even as the VIX Index spikes.

The VVIX Index was introduced in 2012 (CBOE [2012]), but options on the VIX Index begin to be available from 2006, which extends the period for analysis. In the period shortly after introduction of VIX options to the market in 2006, a relatively large truncation error hints at a not-yet-developed market for VIX options.

In general, several patterns emerge from the analysis. For both equity indices, low levels of their volatility indices coincide with higher truncation errors than average. This is consistent for both the SPX Index and the DJIA Index, as well as over time. A possible explanation for this is that the width of the strike price grid with sufficient liquidity shrinks as volatility contracts. This implies that options are becoming cheaper and market participants can afford to purchase protection with higher delta. Another pattern which is consistent for both equity indices is that the truncation error sharply increases during market corrections. This can be explained by the spot price moving beyond the liquid strike grid, which creates large single-sided truncation errors on the left side. The truncation error of the VIX Index behaves completely different. This is fundamentally caused by the different stochastic processes that rule the underlying dynamics.

3.6 Conclusion

This chapter examines the methodology behind the VIX Index and finds that it structurally underestimates MFIV because it ignores the unobservable tails

of the volatility surface. Insights by Lee [2004] into the behaviour of the volatility surface at extreme strikes enable us to derive a compensation term that can correct the structural underestimation in the VIX Index. Historical analysis shows that the underestimation due to truncation fluctuates over time, as option trading activity in the tails varies.

Our approach does not require any assumptions about the underlying process, beyond those laid out by Lee [2004]. The estimation of the extrapolation parameter β for each tail does however add an additional layer of uncertainty to the calculation. It is ultimately the shape of the volatility smile which determines the point where the β-parameters stabilize. Therefore, underlying dynamics and data availability must both be considered carefully when choosing a minimum strike grid width. It is straightforward to extend the suggested approach to incorporate specific model assumptions in scenarios with little available data or long maturities.

Analysis of historical truncation errors has revealed consistent patterns, where the truncation error grows in calm market phases. This implies that volatility indices underestimate forward volatility in calm market phases.

Analysing risk aversion in the market should account for these patterns, as risk appetite of market participants might be overestimated otherwise. The analysis of the term structure of variance can also be improved by this approach. Option chains tend do become less liquid with longer expirations, implying a possible systematic downward bias. The behaviour of the extrapolation factors β_{Left} and β_{Right} as time-to-maturity is increased must therefore be carefully considered.

VIX-like single-stock indices, such as introduced by CBOE [2011], can be based on underlyings with significantly less option trading activity. Compensating for this makes MFIV-estimates comparable between indices and stocks, independently of market conditions. The different market conditions, especially in the single-stock option market, also require further investigation. As strike grid width can be expected to be narrower, the validity of

asymptotic extrapolation on heavily truncated volatility surfaces should be examined closely.

Chapter 4

Martingale Defects in the Volatility Surface and Bubble Conditions in the Underlying

This chapter was published as
Stahl, P. and Blauth, J.: Martingale Defects in the Volatility Surface and Bubble Conditions in the Underlying
Review of Derivatives Research, 27, 85–111 (2024)
https://doi.org/10.1007/s11147-023-09200-x

The martingale theory of bubbles enables testing for asset price bubbles by analyzing option prices. As recently shown by Piiroinen et al. [2018], the SABR model is a strict local martingale when its parameterization implies a positive correlation between stock and option prices. We operationalize this theoretical result and analyze stock price bubbles in 2576 stocks over 26 years. Martingale defect conditions are absorbed quickly by options markets, but identify high proportions in significant and permanent changes in distribution

of price returns, option trading activity, short interest in the underlying, and institutional ownership. These results confirm many common assumptions about stock price bubbles. These bubbles are temporally clustered, and tend to occur in periods of positive market development. Martingale defects are rare in market corrections, which indicates that they are a result of overoptimistic speculation.

4.1 Introduction

An asset price bubble occurs when the price exceeds the fundamental value of the asset for a sustained period. Unfortunately, testing for bubbles suffers from a joint hypothesis problem: As the fundamental value is, by its nature, unknown, both quantities have to be estimated simultaneously, which severely reduces the usefulness of these types of analysis [Camerer, 1989]. The martingale theory of bubbles as developed by Jarrow [1992], Loewenstein and Willard [2000], Cox and Hobson [2005], Heston et al. [2007], Jarrow et al. [2010], Biagini et al. [2014], and many others, characterizes asset price bubbles in terms of strict local martingales.

To overcome the joint hypothesis issue, Jarrow [2015] suggests specifying an option pricing model, which can be validated separately, and checking whether it implies an underlying price process that is a *true martingale*, and not a *strict local martingale*[1]. Piiroinen et al. [2018] derive a martingale defect indicator for SABR dynamics, fundamentally transforming the task of identifying an asset price bubble into a calibration problem. In this chapter, we empirically investigate the connection between stock price bubbles and changes in several non-related variables in a large-scale study of 2576 stocks over approximately 26 years.

We adapt the approach suggested by Piiroinen et al. [2018] to identify mar-

[1]Specifically, see section 3.2.3 of Jarrow [2015]

tingale defects in the volatility surface of single stocks. We evaluate defect persistence, stock price return distribution, option trading activity, short interest, and institutional ownership on a per-event basis and find strong evidence that martingale defects coincide with permanent changes in these variables. Our findings confirm many fundamental assumptions about stock price bubbles. Furthermore, we examine the temporal clustering of martingale defect events over time and find that they predominantly occur during phases of positive market returns and rarely in phases of sustained negative returns.

Monitoring the volatility surface for martingale defects is a useful tool to detect overoptimistic speculation in stocks. It enables investors to identify individual stocks which may not be rationally priced and adjust their exposure accordingly. It is furthermore an appropriate tool for regulators to improve market monitoring and focus their resources in order to protect retail investors. Across the entire market, it provides a gauge of general investor optimism and a markets propensity to develop bubbles.

The remainder of this chapter is structured as follows. To provide background and context to our study, we will begin with a short review of the related literature. Then we examine the details of the martingale theory of bubbles and the implementation by Piiroinen et al. [2018]. Section 4.4 provides an overview of the data, specifics of our calibration procedure, and further implementation details. Our subsequent analysis is two-fold. First, we examine individual bubble events and evaluate changes after observing a martingale defect. Second, we aggregate bubble events across the entire market and investigate their clustering behaviour over time. Section 4.6 concludes the chapter and identifies topics for further exploration.

4.2 Literature Review

The martingale theory of bubbles is fundamentally based on the idea that an asset in bubble conditions is not a fair bet, as market participants are willing to overpay for a potential upside.

Participating in these bubbles can be rational under certain conditions, such as laid out by Jarrow [1992], who analyzes asset price bubbles as market manipulation by large traders and extends arbitrage pricing theory, which is based on an economy of price taking actors, by allowing for large traders. These purchase large positions in individual stocks, and may corner the market or squeeze holders of short positions to force them to pay an arbitrary price. He derives sufficient conditions under which these strategies are impossible under fair pricing measures. In this study, we find that martingale defects in option prices do indeed coincide with increased short interest.

We now review the literature concerning the martingale theory of bubbles, delving into its fundamental assumptions and constraints. In a seminal paper, Delbaen and Schachermayer [1994] generalize the fundamental theorem of asset pricing and introduce *No Free Lunch with Vanishing Risk* (NFLVR), thus paving the way for the martingale theory of bubbles. Jarrow et al. [2007] review the literature on asset price bubbles in complete markets with infinite trading horizons and conclude that, under NFLVR no arbitrage conditions, the existence of bubbles implies that markets must be incomplete. Cheridito et al. [2007] show that it is necessary and sufficient for stocks and bonds to be undominated trading opportunities for an equivalent local martingale measure to exist. Heston et al. [2007] provide conditions on the prices of options to rule out bubbles in the underlying. Jarrow et al. [2010] extend the NFLVR framework by imposing the *no dominance* (ND) conditions suggested by Merton [1973] and allow an infinite number of local martingale measures to coexist, which represent the fundamental economic regimes. For each trade, the market chooses one of these measures to determine the price. When market fundamentals change, a different measure is chosen. While as-

set prices remain unchanged, derivatives prices must change, which implies that stock price bubbles can be inferred from option prices.

Jarrow and Protter [2013] investigate perceived positive alpha under incomplete information, leading to the illusion of arbitrage opportunities in asset price bubbles. Biagini et al. [2020] extend Jarrow and Protter [2013] and study the relationship between private information and perceived bubbles formally, with the potential consequence that asset price bubbles are less likely for assets with less hidden information, such as indices, when compared to single stocks.

The utilization of martingale theory in empirical research proves valuable for examining financial bubbles. Jarrow and Protter [2010] examine the implications for derivatives pricing and detection of asset price bubbles. They suggest three different approaches to detect asset price bubbles. The first approach relies on modelling the fundamental value of assets directly, leading to the joint hypothesis issues described by Camerer [1989]. The second approach specifies a stochastic process for the underlying which is calibrated against the time series of stock price returns. The third approach chooses an option pricing model, and calibrates it against observed option prices. The advantage of this approach is that by calibrating the model, it is automatically validated. In this chapter, we follow the this suggestion by calibrating the SABR model to observed out-of-the-money option prices. Jarrow et al. [2011] employ martingale-based *volatility modelling* to detect bubbles. They estimate a non-parametric volatility function as proposed by Florens-Zmirou [1993] from past asset prices. They illustrate their method based on four stocks during the dotcom-bubble and find that the bubble conditions they identify overlap with stocks and periods that where previously considered bubbles as well. Obayashi et al. [2016] utilize the first approach proposed by Jarrow et al. [2011] to analyze a large number of stocks and examine the lifetime of bubbles. They find that these stock price bubbles exhibit surprisingly long lifetimes, typically spanning several months or years. As irrational exuberance would typically be short-lived, this points towards the rational-

ity of stock price bubbles. In contrast, we find that bubble conditions in the volatility surface are typically absorbed within a few days. In conjunction with our other findings, this lends some support to the notion that bubbles in option markets are predominantly propelled by irrational speculation.

By imposing boundary conditions on option prices at extreme strikes, Jarrow and Kwok [2021] identify bubble conditions in the S&P 500 Index and develop a profitable momentum trading strategy based on the concept of "riding the bubble to the top", as suggested by Conlon [2004], indicating that martingale defects in option prices may have some value for forecasting price returns.

Piiroinen et al. [2018] develop the *martingale defect indicator* from the SABR model by analytically deriving an expression to describe the magnitude of the defect. They analyze the bubble risk of SNAP Inc., Twitter Inc., and Square Inc. in 2017 and 2018 and show that bubble conditions can be accurately detected by calibrating a SABR model to observed option prices. Our empirical investigation relies on this indicator due to its straightforward implementation, as well as the extensive research on the SABR model, which makes its calibration a well-documented process.

Fusari et al. [2022] exploit the put-call price differential to identify bubble conditions from option prices. By estimating a specifically developed *generalized stochastic volatility jump diffusion* or G-SVJD model, which admits both martingale and strict local martingale representations, to put and call options separately, they are able to show that bubbles tend to occur in call options, but not in put options. Furthermore, they find that bubbles tend to occur regularly in single stocks, but rarely in indexes.

A different approach is proposed by Biagini et al. [2022], who suggest training a neural network to recognize whether a smile of call option prices is generated by a strict local martingale or a true martingale. The advantage of this approach is its independence from model specifications, at the cost of computational complexity.

Asset price bubbles are examined not solely through the lens of martingale theory but also via alternative approaches. In our empirical investigation, we are able to substantiate some of these findings. Bakshi et al. [2021] analyze VIX futures curves and find that volatility is mean-reverting and VIX futures are in backwardation when disaster risk is elevated. This provides some empirical backing for the approach of Piiroinen et al. [2018], who identify martingale defects when the correlation between stock price and variance level turns positive.

The concept of *dark matter* in asset pricing models draws a connection to economic components that are difficult to measure directly and quantifies its impact on model stability, which was formalized by Chen et al. [2022]. Using a semimartingale-based approach, Bakshi et al. [2022] show how to decompose equity option risk premiums and examine the dynamics of *jumps crossing the strike* and *local time*. Intuitively, our approach identifies bubble conditions in the underlying when the correlation of price and volatility is positive, which suggests a possible link with the findings of Bakshi et al. [2022], who show that negative premia for upside equity risk are consistent with the presence of unspanned risks.

Blocher et al. [2021] find that short sellers are regularly forced to exit positions earlier than optimal. In our study, we find that bubble conditions induce increased short selling activity, supporting their short squeeze hypothesis. SEC [2021] examine market structure and trading activity around the GameStop Inc. bubble in January 2021 and find, contrary to common perception at that time, that short covering was not the primary driver of the stock price run-up. Instead, the primary driver for this specific bubble event were overly optimistic, young, inexperienced investors. Observing overoptimistic speculation in the volatility surface further corroborates these findings. Mohrschladt and Schneider [2021] link option prices with internet search interest and find that retail investors contribute to idiosyncratic volatility through irrational trades, which are exploited by sophisticated market participants in the options market.

4.3 Theoretical Background

In this section we describe the market setting and adapt the approach by Piiroinen et al. [2019] to utilize the SABR model to detect bubble conditions for a large data set.

We fix a finite time horizon $T > 0$ and consider the filtered probability space $(\Omega, F, (F_t)_{t\geq 0}, \mathbb{Q})$, where F is the σ-field of measurable subsets of Ω and the filtration $(F_t)_{t\geq 0}$ satisfies the usual conditions (see Protter [2016]). $\mathbb{Q}$ denotes an equivalent local martingale measure and, by possibly embedding $(\Omega, F, (F_t)_{t\geq 0})$ into a larger complete market, we can assume that $\mathbb{Q}$ is also unique. We consider a stock price process $(S_t)_{t\geq 0}$ on Ω with continuous paths $\mathbb{Q}$ almost surely. Its forward price is then given by

$$F_t = S_t e^{(r-d)t},$$

for some constant risk free rate r and dividend yield d which might include borrow costs. We denote the expectation operator with respect to $\mathbb{Q}$ by $\mathbb{E}$ and, if we want to empathize on the current price $x = S_0 \geq 0$, we put x into the subscript $\mathbb{E}_x$.

We follow Piiroinen et al. [2019] by making the following

Definition 1 *The stock price S_t is said to admit a bubble on $[0, T]$ with respect to $\mathbb{Q}$ if the discounted process*

$$S_t e^{-(r-d)t}$$

is a strict local martingale on $[0, T]$ with respect to $\mathbb{Q}$. The normalized martingale defect is defined as

$$d_x(T) := 1 - \frac{e^{(d-r)T}}{x}\mathbb{E}_x[S_T].$$

The intuition behind this definition is as follows: the asset is currently in the state of a bubble if the current market value of the asset S_0 exceeds its fundamental value which is the discounted expectation of its future value $e^{-rT}\,\mathbb{E}[S_T]$.

From the definition of a (local) martingale it is immediate that our martingale defect indicator satisfies $d_x(T) \geq 0$ with equality if and only if our model does not admit a bubble.

It shall be noted that d_x not only tells us whether or not an asset is currently in bubble condition but also, due to its domain $d_x \in [0, 1]$, quantifies the strength of the bubble: the closer d_x is to one, the stronger the indication of a bubble event.

We use the results in Piiroinen et al. [2019] to calculate an analytical expression for d_x from market data. More precisely, we use available option prices for a given asset, calibrate a SABR model to this data and use the resulting parameters to compute our bubble indicator. This reduces our problem of estimating d_x to the problem of calibrating a volatility smile and leads to a particularly simple representation of d_x in the model parameters. We start by giving a rough overview over the model.

The stochastic volatility model introduced by Hagan et al. [2002], commonly known as *stochastic alpha, beta, rho* or SABR model, is an extension of the CEV model and is determined by the SDEs

$$dF_t = \alpha_t F_t^{\beta} dZ_t^1 \tag{4.1}$$

$$d\alpha_t = \nu \alpha_t dZ_t^2, \tag{4.2}$$

where $\alpha_0 = \alpha$ and

$$\langle dZ_t^1 dZ_t^2 \rangle = \rho t. \tag{4.3}$$

The elasticity parameter $\beta \in [0, 1]$ controls the general behaviour of the

model and is usually not calibrated but chosen in advance. Following Piiroinen et al. [2019, Theorem 3.1], we set $\beta = 1$, which is the log-normal case. $\nu \geq 0$ is the volatility-of-volatility and $\rho \in [-1, 1]$ is the correlation between the driving Brownian motions Z^1 and Z^2.

The advantage of the SABR model in this context is that it admits a strict local martingale representation, unlike many traditionally used option pricing models. The SABR model therefore implies that stock price bubbles can possibly exist. Fusari et al. [2022] examine this issue in further detail, and contribute a more sophisticated approach based on their G-SVJD model.

For notational clarity and without loss of generality, from now on, interest rates and dividend are omitted and assumed to be zero. Piiroinen et al. [2019] then show that under SABR dynamics the martingale defect takes the particularly simple form

$$A(\alpha, \nu, \rho) := \lim_{T \to \infty} d_x(T) = 1 - e^{\frac{-2\rho\alpha}{\nu}}. \tag{4.4}$$

This is Piiroinen et al. [2018, Equation 15].

Since both volatility α and volatility-of-volatility ν must be strictly positive, A can only become positive when the correlation between stock price and volatility ρ is positive as well. Intuitively, this means that options get more expensive as the stock experiences positive returns, hinting at optimistic risk-seeking preferences.

4.4 Data and Calibration

Our analysis covers the constituents of the MSCI USA Investible Market Index, which attempts to cover all large-, mid- and small-cap US stocks[2]. We analyze daily close prices of all put and call options on all stocks within

[2]MSCI [2017] documents the index' construction.

the universe within a set of restrictions outlined below. In total, 2576 stocks are considered, and our universe is updated monthly to match the index. Constituent information is provided by MSCI. The option data is provided by OptionMetrics, begins on January 1st, 1996, and ends on April 25th, 2022, spanning 6758 trading days. The options dataset consists of daily close prices for approximately 51.3 million option contracts on all analyzed stocks. Stocks for which no options have been traded are excluded from the analysis. The option price data is matched to price data of the underlying price and to linearly interpolated U.S. treasury rates. Dividends are assumed to be a constant yield, based on the last known dividend payment. American options are evaluated using the approach suggested by Cox et al. [1979]. Average daily trading volume, outstanding short interest, and institutional holdings data have been provided by Bloomberg.

Options with a daily trading volume of less than 100 contracts are excluded for that day. We write $k_i = log\left(\frac{K_i}{F_T}\right)$ for the log moneyness. Our calibration only considers out-of-the-money options, as they tend to be more liquid than in-the-money options. This entails using call implied volatilities for strikes below the forward price and put implied volatilities for strikes exceeding the forward price. As shown by Fusari et al. [2022], we will not be able to see a bubble solely based on the put options. However, it is worth highlighting that the martingale defect indicator, as it is fundamentally derived from the parameter ρ, effectively models the skew of the volatility smile. In the SABR model, this parameter characterizes the 'difference' between the call and put wings of the smile. In this regard, our approach is consistent with Fusari et al. [2022]. We further restrict the calibration to options with absolute log-moneyness $|k| < 0.5$.

This leaves us with a number of pairs $(K_i, \sigma_i)_i$ of strikes and implied volatilities for the remaining call or put options. The remaining contracts are enumerated by running variable $i = 1, \ldots, N$, where N is the total number of remaining contracts. If $N < 5$, the stock is excluded from the analysis for that day. On average, each calibration is fitted against 6.12 contracts. This

number increases over time as more contracts are traded. Figure B.1 in the appendix provides the average number of available contracts per calibration over time.

As the volatility process 4.2 is not mean reverting, the SABR model is better suited for short expirations [Gatheral, 2006, Ch. 7]. Our analysis is therefore two-fold. We calibrate the model against the full option price surface and refer to these results as *full surface*. We repeat the analysis, but restrict the calibration of the model to options with one month to expiration. For the *one-month-tenor*, contracts are selected to have a remaining lifetime between 25 and 35 days. We do not use shorter expirations, because the relative volume has shifted within the time frame of our analysis. Options with short maturities became more popular within the time frame of our analysis[3].

We calibrate the model to the remaining option mid prices, which we convert to implied volatilities. Our minimization problem is

$$\underset{\alpha,\nu,\rho}{\text{minimize}} \sum_{i=1}^{n} \left(\sigma_{\alpha,\nu,\rho}(k_i) - \sigma_i\right)^2 \tag{4.5}$$

s.t.

$$\alpha \geq 0 \tag{4.6}$$

$$\nu > 0 \tag{4.7}$$

$$\rho \in (-1, 1) \tag{4.8}$$

where $\sigma_{\alpha,\nu,\rho}(\cdot)$ is the SABR implied volatility function for given parameters α, ν, ρ and fixed time to maturity T.

We implement the suggestion of Le Floc'h and Kennedy [2014], who present an explicit initial guess procedure to generate an initial parameter guess for the minimization problem. This initial guess is the starting point for a Nelder and Mead [1965]-minimization of equation 4.5. We repeat this procedure for

[3]Appendix B.3 shows that the relative proportion of contracts shifts to the short end from 2012 onward.

both the one-month tenor as well as the full surface calibration daily for every stock under consideration. In order to avoid redundancy, we occasionally present only results for the full surface in the text body and move one-month tenor to appendix B.5, where appropriate.

To validate the model calibration, we calculate the at-the-money SABR-model implied volatility of the root mean squared error on a 30 day horizon $RMSE_{IV}$. In total, we have 1.4 million calibrations with a mean $RMSE_{IV}$ of 0.97%. Of these, 56722 calibrations (4.05% of all calibrations) show a martingale defect $A(\cdot) \geq 0.05$, with a mean $RMSE_{IV}$ of 1.99%. Appendix B.1 provides full results.

From these calibrated parameters α, ν and ρ, the martingale defect $A(\cdot)$ follows from equation 4.4. Due to the large number of stocks in our study and the computational effort required for Markov Chain Monte Carlo methods, we forgo obtaining a distribution for the martingale defect as suggested by Piiroinen et al. [2019].

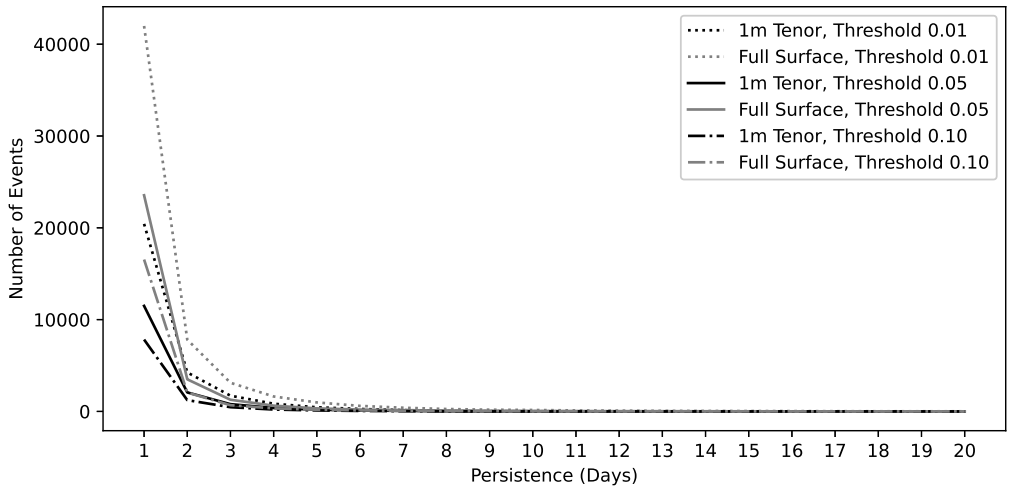

Figure 4.1: Relative frequencies of persistence of martingale defect events in trading days. Martingale defects based on SABR calibration of the 1-month tenor as well as the full volatility surface with thresholds $A_{min} = \{0.01, 0.05, 0.1\}$.

For our analysis across a large number of stocks we must convert the

martingale defect indicator $A(\alpha, \nu, \rho)$ of equation 4.5 into a binary signal $I(A(\cdot)) \in \{0, 1\}$. A positive signal marks an event in our study, which is specific to a single stock and date, at which the option price surface for this stock indicates bubble conditions for its underlying.

By design, the defect indicator be will greater when the volatility surface deformation is more pronounced, and very small values indicate a low bubble intensity [Piiroinen et al., 2018]. To reduce the noise created by the faintest bubble conditions, we impose a minimum threshold A_{min} on the indicator[4]. Since the martingale defect indicator is based on the correlation ρ, which is normalized and thus time independent, we pick a static threshold A_{min}.

Asset price bubbles are in part defined by their persistence. We utilize this property to fortify our analysis against calibration issues which may not have been dealt with by the calibration procedure described in section 4.4. By requiring $A(\cdot)$ to remain above a threshold for a number of consecutive days p, the total number of events is reduced, but the accuracy of the signal can be improved. Section 4.5.1 examines defect persistence in more detail. Days where we cannot find a sufficiently accurate calibration are excluded, and break persistence. Our indicator is thus defined as

$$I_t = \begin{cases} 1, \text{ if } A_{t-n} \geq A_{min} \forall\ n \in [0, \ldots, p] \\ 0, \text{ otherwise.} \end{cases} \tag{4.9}$$

To prevent counting the same fundamental event multiple times, event periods must be non-overlapping. Where SABR model is calibrated against the full surface, the indicator is denoted I^{full}, and where it is calibrated against the one-month tenor of the volatility surface, it is denoted I^{1m}. Figure 4.1 provides the numbers of events for various thresholds and persistence requirements. Increasing these two requirements reduces the total number of events considered as well as the number of affected companies. We will further ex-

[4]Appendix B.1 provides summary statistics of all events conditional on the martingale defect level.

amine this in the next section. Calibrating the SABR model against the full surface generates a larger number of bubble events because the number of datapoints for each calibration is larger, and calibration errors are less likely. Consequently, persistent events are less likely to be interrupted by miscalibrations. Since calibrating the SABR model on the full volatility surface tends to smooth out anomalies in the low-DTE part, we expect to see more false negatives (*Type II errors*) compared to the one-month tenor.

Most events do not persist over multiple days, which will be examined in detail in the next section. While we observe 18526 events with a persistence of one day above a threshold of $A_{min} = 0.01$ based on the one-month tenor, this count reduces to 3505 after two days, and to 1339 after three days. Based on the full surface, we observe 38798 events with a persistence of one day, which reduces to 6731 after two, and 2477 after three days. The effect of A_{min} is similar. For $A_{min} = 0.05$, only 9983 events remain for the one-month tenor, while 21087 events remain on the full surface calibration. The effect is even stronger for $A_{min} = 0.1$, with 6659 and 14602 events remaining before applying a persistence requirement. A threshold of $A_{min} = 0.05$ with a minimum persistence of 2 trading days appears to strike a balance between sensitivity and noise, and are chosen for the remaining analysis. With these requirements, 1518 events remain for the one-month tenor, and 2745 events remain when calibrating against the full surface.

4.5 Analysis

In order to investigate the relationship between martingale defects derived from the volatility surface and bubble conditions in the underlying, we assess return distribution and trading activity before and after detecting a martingale defect event. Furthermore, we examine the relative occurrence of martingale defects over time.

For each event, we consider data during a certain number of trading days before and after the event, which we refer to as *event period* t_{event}. To assess whether an effect is persistent, we provide results for multiple event periods

$$t_{event} \in \{21, 63, 84, 105, 126, 189, 252\}.$$

Since bubbles are hard to quantify, we analyze a variety of metrics to establish a tight-knit connection between martingale defect and bubble conditions in the underlying. Since it is the foundation of our analysis, we begin with the reaction of the option market to a martingale defect event. Assessing the persistence of martingale defects in the volatility surface gauges the ability of the options market to arbitrage irregularities away. Next, we analyze whether the distribution of returns changes after observing a defect. Using the number of actively traded option contracts, we find that the martingale defect coincides with increasing option trading activity for this underlying. We also find that outstanding short interest increases. By examining institutional ownership, we find that institutional traders tend to reduce their exposure to stocks in suspected bubble conditions, leaving the participation to retail investors.

These four characteristics are naturally time-varying, and might exhibit time dependent variation that are unrelated to stock price bubbles. To avoid confounding the bubble-induced change with natural variation over time, we study a set of placebo-events and compare the results to identified bubble events. We generate these placebo events on a per-stock basis by shuffling the dates at which martingale defect events have been identified. This way, the number of events per stock remains identical, but the timing of the events is randomized. The events are restricted to the time period where each stock is a constituent of the IMI Index. We repeat the analysis on this set of placebo events and report the results for comparison.

Finally, we investigate bubble events over time across the entire market and find that martingale defects tend to occur more often in good times, and

rarely in bad times.

4.5.1 Defect Persistence

The martingale defect fundamentally indicates an irrational deformation of the volatility surface.

Figure 4.1 shows the frequency of event persistence for all incidences with $A_{min} \in \{0.01, 0.05, 0.1\}$. The majority of events occurs for only one day, and frequency drops quickly. The longest event in the one-month tenor analysis was Sundial Growers Inc., where a defect was indicated for 14 consecutive days in May 2021. In the full surface analysis, Myriad Genetics Inc. indicated a defect for 20 consecutive days in December 2007. PubMatic Inc. also indicated a 20 day defect in June 2021. All three companies are well-known as *meme stocks* on various retail trading investing websites[5].

The short persistence of the martingale defect hints at the efficiency of options markets to absorb bubble conditions in the underlying into a rational shape of the volatility surface. It does not imply that irrational exuberance in the underlying is not persistent, but rather that options markets can accommodate and return to efficiency quickly. Obayashi et al. [2016] analyze the lifetime of financial bubbles by modelling the distribution of the underlying directly, and find that bubble conditions in the underlying persist on the scale of years.

Appendix B.4 presents more details on the effect of threshold and persistence on the number of observed events, as well as the number of affected companies.

[5]For example, see Sun [2021], Smith [2007], Farooque [2022]

4.5.2 Change of Distribution

To confirm the connection between martingale defect events in the volatility surface and the price process of the underlying, we analyze whether the distribution of historical log returns changes with an event.

To assess whether the martingale defect indicator can reliably identify a change in the distribution of log-returns of the underlying, we employ the two-sample *Kolmogorov-Smirnov (K-S)*-test. The Null-hypothesis is that the distribution of log-returns during a period of given length before and after an event is identical. The interpretation of this would be that observing a martingale defect in the volatility surface does not coincide with a changing log-return distribution of the underlying, and is purely an anomaly in the volatility surface.

More specifically, for given underlying and length of event period $t_{event} > 0$ and for each date t_0 where we observe a martingale defect signal, we perform a K-S-test between the two sets $\{r_{t_0-1}, \dots r_{t_0-t_{event}}\}$ and $\{r_{t_0+1}, \dots r_{t_0+t_{event}}\}$, where we denote by

$$r_i = log\left(\frac{S_{t_i}}{S_{t_i-1}}\right) \tag{4.10}$$

the return on day t_i.

For each event, we compute the K-S test statistic and p-value. Finally, we calculate the 1%-quantile and 5%-quantile of all p-values, which we report for a range of event periods in table 4.1. These quantiles represent the proportion of events where the Null hypothesis that distribution of log-returns before and after an event is equal can be rejected with a confidence level of 1% and 5% respectively. By comparing different time frames, we gain insight into the persistence of changes. It should be emphasized that, regardless of the event period length, the martingale defect events are identical. The number of events changes only where event periods overlap or extend beyond the

	One-Month Tenor					Full Surface				
Period t_{event}	N	$p \leq 1\%$		$p \leq 5\%$		N	$p \leq 1\%$		$p \leq 5\%$	
63	867	12.31%	(1.29%)	22.99%	(3.28%)	1402	11.40%	(0.94%)	21.65%	(2.18%)
84	860	17.25%	(1.86%)	28.80%	(5.71%)	1393	15.24%	(2.50%)	26.44%	(6.82%)
105	836	21.09%	(2.31%)	33.81%	(8.29%)	1371	19.97%	(2.44%)	31.94%	(10.26%)
126	820	23.93%	(4.29%)	37.00%	(14.54%)	1345	23.56%	(3.50%)	36.61%	(11.71%)
189	793	36.69%	(4.07%)	50.55%	(15.03%)	1310	35.14%	(3.81%)	48.58%	(13.36%)
252	739	47.07%	(10.57%)	59.66%	(22.05%)	1245	47.71%	(8.47%)	59.20%	(20.06%)

Table 4.1: Aggregate results of two-sample Kolmogorov-Smirnov tests to compare empirical daily return distributions of the underlying before and after an event. Results on the left side are based on the calibration of the SABR model against the one-month tenor of the volatility surface, while results on the right side are based on the full surface. Event Period t_{event} provides the number of trading days before and after an event under consideration. N refers to the number of events after discarding stocks with insufficient data and overlapping event periods. $p \leq 5\%$ and $p \leq 1\%$ report the quantile of p-values in the sample. The corresponding results of the placebo study with shuffled event dates are shown in parentheses.

available data.

We observe that, for both the 1%-quantile and 5%-quantile, a longer event period increases the proportion of significantly different return distributions. For an event period of 252 trading days, almost 60% of events reject the null hypothesis of an unchanged log-return distribution with a confidence of 95%. Even though the options market absorbs the martingale defect within a few days, these findings imply that it hints at a permanent change of the underlying's price process for a large proportion of events. A longer event period increases the number of datapoints and reduces uncertainty, therefore longer event periods are inherently more reliable. It is therefore not clear whether the effect is immediate or takes some time to manifest.

In our placebo study, we shuffle the identified event dates for each stock in order to randomize the timing component. The randomized sample shows a highly reduced proportion of significant changes in the return distribution, indicating that this distribution is indeed changing over time but to a generally lesser degree than when only accounting for bubble events.

4.5.3 Option Trading Activity

Period t_{event}	Full Surface						One-month Tenor					
	N	t-statistics	p-values				N	t-statistics	p-values			
		Mean	$p \leq 5\%$		$p \leq 1\%$			Mean	$p \leq 5\%$		$p \leq 1\%$	
21	780	0.16	27.77%	(8.08%)	19.20%	(3.35%)	670	0.36	13.75%	(3.58%)	6.36%	(1.15%)
63	800	0.47	41.52%	(15.34%)	32.76%	(9.72%)	724	0.33	26.07%	(6.79%)	16.03%	(2.69%)
84	797	0.57	43.26%	(25.34%)	35.00%	(19.75%)	702	0.42	25.00%	(14.16%)	13.84%	(8.32%)
105	775	0.82	46.47%	(25.27%)	38.68%	(18.48%)	692	0.54	28.14%	(12.36%)	17.84%	(8.29%)
126	755	0.94	49.15%	(26.18%)	40.85%	(22.44%)	673	0.62	33.17%	(17.17%)	21.83%	(9.97%)
189	735	1.39	53.41%	(30.48%)	44.82%	(24.44%)	660	0.90	36.99%	(16.57%)	28.16%	(12.08%)
252	677	2.03	55.21%	(32.16%)	48.85%	(26.49%)	611	1.12	41.81%	(20.06%)	31.94%	(14.40%)

Table 4.2: Aggregate results for t-tests of the number of actively traded option contracts before and after a martingale defect event, based on the calibration of the SABR model against the one-month tenor of the volatility surface. Event Period t_{event} provides the number of trading days before and after an event under consideration. Results under *Full Surface* count all available options, while results under *One-month Tenor* count only contracts within the one-month tenor. N refers to the number of events after discarding stocks with insufficient data and overlapping event periods. Mean t-statistic reports the mean of the t-statistic of all events. $p \leq 5\%$ and $p \leq 1\%$ report the quantile of p-values in the sample. The corresponding results of the placebo study with shuffled event dates are shown in parentheses.

In order to examine the relationship between a martingale defect in the volatility surface and option trading activity, we analyze whether the number of contracts available remains constant before and after an event.

For each day, we calculate the number of traded options contracts for the underlying, while applying the liquidity requirements laid out in section 4.4. For a pre-specified event window length, we calculate the average number of daily available contracts before and after each event. We calculate results counting contracts on either the full surface or restrict ourselves to the one-month tenor. This is independent of the tenor selection for the calibration. Over the time period of our sample, the number of daily actively traded options has grown considerably. Appendix B.2 provides a short overview of option trading activity over time. Therefore, the number of active contracts after an event should be expected to be larger than before, just due to the length of the event period. To compensate for this positive trend in the number of available options we adjust the observations downwards by cal-

culating the total growth in options contracts during estimation and event period and distributing it equally across stocks. This implies (falsely) that option trading activity grows at a constant rate during the event period, and is equally distributed amongst securities. The effect of this adjustment is on average 0.0014 contracts for an event period of 10 days, and 0.0354 contracts for an event period of one year, and is negligible. The remaining difference can therefore be attributed to the martingale defect event. The number of contracts is consistently higher after an event. We evaluate this effect using a paired t-test, i.e. we compare the number of options contracts before and after an event on a per-event basis. Results are reported in table 4.2. The resulting p-values are consistently very small across event periods, indicating that the increased number of traded contracts is significant. The mean difference grows with longer event windows and does not rebound after a while. This implies that the cause for the increase is not a sudden price movement, but rather a persistent effect created by increased options trading activity for the affected underlying. The observed effect is much stronger when considering all available options, not only those within the one-month tenor, implying that the market prefers either shorter or longer options for bubble speculation. As the number of daily available contracts has grown continuously over time, the placebo study shows a similar pattern over time, but for a much smaller proportion of events.

4.5.4 Short Interest Ratio

Jarrow [1992] develops the martingale defect theory of bubbles to investigate price manipulation by large traders. One of his examples involves cornering the market and squeezing the holders of short positions to pay any price arbitrarily chosen by the large trader[6]. He provides two reasons why this might happen. First, short traders are unable to observe the large traders purchases, thus not realize that the market is cornered. Second, that the

[6]Specifically, example 2 in Jarrow [1992].

Period t_{event}	N	t-statistics	p-values			
		Mean	$p \leq 5\%$		$p \leq 1\%$	
21	683	1.454	75.3%	(51.7%)	66.3%	(47.0%)
63	670	1.331	79.6%	(59.4%)	74.0%	(61.0%)
84	661	1.069	80.8%	(57.9%)	75.2%	(55.7%)
105	635	0.614	81.6%	(70.6%)	76.9%	(62.1%)
126	623	0.260	82.7%	(74.8%)	78.2%	(63.9%)
189	594	0.336	85.2%	(70.9%)	80.1%	(67.1%)
252	538	0.528	86.6%	(78.2%)	82.1%	(74.2%)

Table 4.3: Aggregate results for t-tests of outstanding short interest divided by ADTV before and after a martingale defect event, based on the calibration of the SABR model against the one-month tenor of the volatility surface. Event Period t_{event} provides the number of trading days before and after an event under consideration. N refers to the number of events after discarding stocks with insufficient data and overlapping event periods. Mean t-statistic reports the mean of the t-statistic of all events. $p \leq 5\%$ and $p \leq 1\%$ report the quantile of p-values in the sample. The corresponding results of the placebo study with shuffled event dates are shown in parentheses.

cornering is technical in nature, i.e. that the short position exceeds the floating supply of shares, and the large traders position exceeds the float.

To investigate this hypothesis, we analyze the short interest in a stock relative to its available supply before and after observing a bubble event. As a proxy for freely available shares we use *Average Daily Trading Volume (ADTV)* as reported by the exchange. We divide outstanding short interest by ADTV, which is commonly referred to as *days-to-cover ratio*, as it measures the number of days would take short sellers buying the entire trading volume to cover their short positions.

For each stock, we retrieve daily outstanding short interest and ADTV, and calculate the days-to-cover ratio. Stocks where either is missing are removed from the sample. Individual missing values are filled forward. For each event, we calculate t-statistics of the days-to-cover ratio before and after the event over a certain number of trading days in the same fashion as above. We repeat this analysis for multiple event periods from 21 to 252 days. The

results of the t-tests are aggregated.

Table 4.3 provides results based on the events generated by the I^{1m}. The mean t-statistic is positive for all event periods, with the largest value being 2.185 for the shortest event period, and ranging between 0.647 and 0.998 for event periods longer than three months. This implies that the short interest rises after a martingale defect is observed in the volatility smile. It rises, on average, by roughly one standard deviation of the days-to-cover ratio. For an event period of 21 days, 75.2% of events have a p-value of less than 5%, and 66.7% have a p-value of less than 1%. With longer event periods, and more daily observations per event, the proportion of significant test results rises. The longest event period shows the largest proportion of highly significant results, with 88.0% of events have a p-value of less than 5%, and 83.7% have a p-value of less than 1%. Appendix B.5 provides similar results based on I^{1m} in table B.5.

After a martingale defect in the volatility surface, the market appears to be increasing short positions. For a very large proportion of events, this is highly significant. The effect is largest in the short term and levels off after a few months, but remains elevated for at least a year. This seems to confirm that the martingale defect can reveal bubble conditions in the underlying, and that some traders have similar perceptions. Similar to the previous results, the placebo study shows a similar pattern of growth with longer event periods, but with significantly reduced proportions. While short interest seems to be on an upward trend long event periods, martingale defects provide a clear indication of distinct increases in short interest.

4.5.5 Institutional Ownership

In recent years, retail traders have shown increased interest in option markets [Deshpande et al., 2020]. We investigate whether martingale defects spike the interest of retail investors by analyzing the percentage of shares held by

Period t_{event}	N	t-statistics	p-values			
		Mean	$p \leq 5\%$		$p \leq 1\%$	
21	880	-0.585	62.0%	(43.1%)	54.5%	(34.2%)
63	867	-1.031	68.1%	(49.1%)	64.8%	(43.8%)
84	860	-0.193	67.4%	(48.6%)	65.3%	(49.4%)
105	836	-0.178	67.3%	(57.4%)	65.4%	(54.0%)
126	820	-0.671	67.4%	(58.5%)	65.9%	(56.4%)
189	793	-0.326	66.5%	(57.4%)	64.5%	(53.5%)
252	739	-0.088	66.9%	(59.6%)	65.3%	(56.8%)

Table 4.4: Aggregate results for t-tests of the percentage of institutional ownership before and after a martingale defect event, based on the calibration of the SABR model against the one-month tenor of the volatility surface. Event Period t_{event} provides the number of trading days before and after an event under consideration. N refers to the number of events after discarding stocks with insufficient data and overlapping event periods. Mean t-statistic reports the mean of the t-statistic of all events. $p \leq 5\%$ and $p \leq 1\%$ report the quantile of p-values in the sample. The corresponding results of the placebo study with shuffled event dates are shown in parentheses.

institutions before and after an event.

Our data is matched to institutional holdings data provided by Bloomberg, which includes the holdings of institutions of type 13F, US and International Mutual Funds, US Insurance Companies, and aggregate institutional stake holdings[7]. This data is available on a weekly basis since March 2010. Stocks where either is missing are removed from the sample. Individual missing values are filled forward. For each event, we calculate t-statistics of institutional ownership before and after the event over a certain number of trading days. We repeat this analysis for multiple event periods from 21 to 252 days. The results of the t-tests are aggregated.

Table B.6 provides results based on the events generated by the I^{full}. The mean t-statistic is negative for all event periods, however, no clear trend is apparent. It ranges from -0.102 to -0.479. In a similar pattern as before, with longer event periods, and more daily observations per event, the pro-

[7]For further information, see Bloomberg Terminal `FLDS DS211`.

portion of significant test results rises. The longest event period shows the largest proportion of highly significant results, with 64.7% of events have a p-value of less than 5%, and 63.1% have a p-value of less than 1%. Appendix B.5 provides similar results based on I^{1m} in table 4.4.

We find strong evidence that after a martingale defect in the volatility surface the percentage of institutional ownership tends to be lower than before. Our analysis does not reveal whether this is due to institutional investors reducing their position in response to bubble conditions, or due to increased retail trader demand, and is limited to ownership of the stock, not the options. In addition, the placebo study reveals an overarching trend of slowly increasing non-institutional market participation. Proportions of significant changes are lower throughout, with a larger gap for shorter event periods, indicating that martingale defects mark distinct short-term changes in institutional ownership even under long-term trends.

4.5.6 Bubbles as Market-wide phenomenon

We will now analyze whether the occurrence of martingale defects between stocks are related to each other by examining temporal clustering of events. We show that bubbles happen predominantly in good times, and occurrence of bubbles falls sharply when the overall market corrects.

For our entire stock universe, we count daily bubble events. During the time span of our analysis, option trading activity has grown[8]. The number of observed events grows in line with this trend. This does not indicate that more assets are in bubble conditions, rather that these conditions are now tradeable in the options market. Figure 4.2 reports the total count of daily events as a fraction of daily total liquid contracts with the intention of compensating this development for our analysis. For the one-month tenor, we use the total number of traded contracts within that tenor as denominator. Since the

[8]Appendix B.2 examines the number of daily liquid option contracts over time.

number of active contracts fluctuates daily, we use 252 trading day rolling averages for normalization. As a proxy for overall market performance, we use the S&P 500 Index.

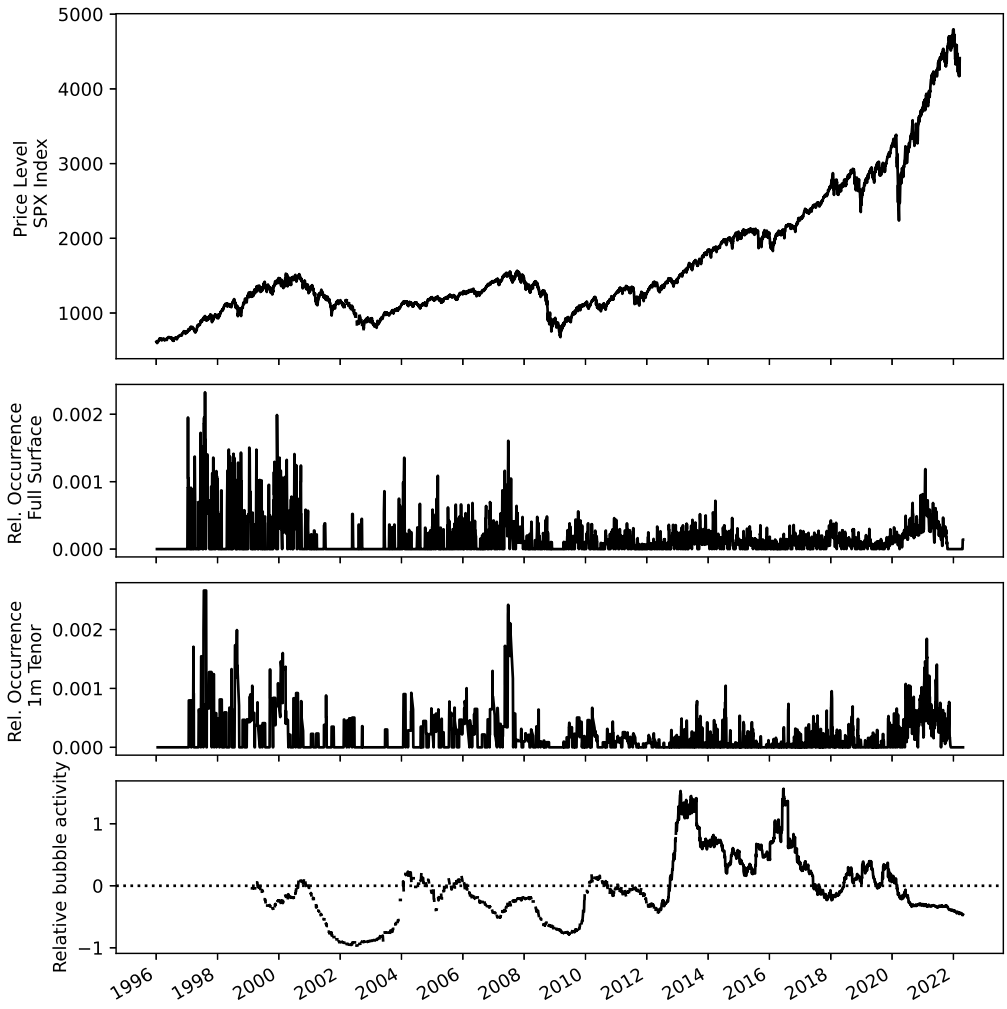

Figure 4.2: Number of daily bubble events divided by the number of daily actively traded options contracts for the entire stock universe. Top panel shows the performance of the S&P 500 Index. Second panel shows the relative occurrence of bubble events based on the full surface. Third panel shows the relative occurrence of bubbles based on the one-month tenor. Bottom panel compares relative occurrences from panel two and three, smoothed by 252 trading days.

Relative occurrence of bubbles based on both the full surface and the one-month tenor follow a very similar pattern. Over time, the relative number

of bubble events fluctuates between zero and 0.25% of daily liquid options. From 1997 to 2000, the SPX Index had positive performance, and the relative occurrence of martingale defects was highest in the entire sample. In the recession of 2001, the relative number of martingale defects was lower and less clustered, showing waning optimism. After a short rebound, the number of bubble events increases and is more clustered than during the drawdown. A further market correction until June 2003 disappoints the optimism, and a relatively bubble-free period commences until January 2004. As market performance is positive, the relative number of bubbles increases steadily, reaching a peak at the market top in 2007. During the recession, the number of bubble events falls quickly, and there are no bubble events between November 2008 and February 2009. From 2009 to 2012, SPX Index performance is positive, and the number of martingale defects grows. As a reaction to the 2012 correction, the number of events collapses, but continues to be elevated until 2016, after which it slightly levels off. At the beginning of 2020, the number of bubble events reaches 0.5%, which it has not reached since 2014. During the Covid-19 market correction, the number of events drops off sharply but rises quickly afterwards and remains elevated throughout the pandemic.

Overall, we observe that long periods of positive returns lead to a rise in martingale defect events in single stock volatility surfaces. By construction, our indicator is elevated when – in the parameters of the SABR model – stock price and volatility become positively correlated. Our observation implies that single stock options are used by the market to express strong optimism, and that option prices are higher than rational. These results support the hypothesis that asset price bubbles are an expression of overly optimistic expectations.

From 2016 on, the number of actively traded contracts begins to rise at a larger rate than before. The number of martingale defect events in relation to the total number of contracts is only very slightly lower. This implies that the multitude of single stocks with active option markets now absorb

the gambling needs of the market. The bottom panel of figure 4.2 compares relative occurrence of bubble events of the one-month tenor to those of the entire surface. A negative value means that bubble events are more prevalent in the one-month tenor, while a positive value means the opposite. While monthly contracts appear to be preferred for trading bubble events for the majority of our sample, trading activity shifted with the introduction of weekly and bi-weekly options. The expansion of options markets into shorter expirations and smaller companies does not appear to have created more bubble events, but redistributed them among a larger number of smaller securities.

4.6 Conclusion

This chapter provides a large scale study of martingale defects in the volatility surface and changes in stock price dynamics and trading reaction. We find that martingale defects tend to coincide with other bubble characteristics.

We operationalize the detection of stock price bubbles by calibrating a SABR model to observed option prices, simplifying the approach suggested by Piiroinen et al. [2018]. Using a large stock and option price dataset, we calculate a daily bubble indicator for the constituents of the MSCI IMI Index.

The volatility surface admits bubble conditions regularly, but rarely for longer than three days, which implies that the options market is generally efficient absorbing bubble conditions.

For each identified martingale defect event, we analyze changes of the return distribution, option trading activity, short interest, and institutional ownership. For all four factors, we find that a large proportion changes significantly after bubble conditions are observed. These effects appear to be permanent. By comparing our results with those obtained from a set of event dates that have been randomly shuffled, we affirm the reliability of martingale defects

in identifying bubble conditions. This shows that our methodology successfully avoids any potential confounding of stock price bubbles with naturally changing characteristics.

The empirical distribution of returns is significantly different for almost 60% of events after one year. Option trading activity, measured by actively traded contracts of any maturity, remains significantly elevated a year after an event for about 55% of events. Short interest remains significantly increased after a year for 88% of events. Institutional ownership decreases significantly within the year after 67% of events. Our results become more robust as the event period increases, and all examined effects appear to be permanent.

Our market-wide analysis of bubble events over time reveals that martingale defects in the volatility surface tend to occur in periods of positive market returns. Market corrections lead to an immediate collapse of the number of martingale defects across all stocks. As option contract availability increased over time, martingale defects are distributed across more stocks, but the relative occurrence of defects remains somewhat constant. The implication is that overoptimistic speculation is a constant element of markets, and that better availability of options can help distribute this irrational force across a larger number of underlyings.

The advantages of the presented simplified implementation of the martingale defect indicator make it a promising tool for future research. Of particular interest might be the correlation between the analyzed factors, the distribution of option trading activity between retail traders and institutional investors to isolate the propensity to take more risk than intended, and the particular effect of options with a lifetime of less than two weeks. Furthermore, our results might be of interest to regulators to help balance option availability, since an overabundance of options for small stocks might permit easier market cornering and thus manipulation, negatively affecting market stability and trust in institutions.

Chapter 5

Concluding Remarks

> *Options and option pricing theory give us an opportunity to measure directly the degree of anticipated uncertainty in the markets.*
> [Ross, 1987]

The volatility surface offers unique insights into the markets perception of risk. It is the subject of an extensive body of literature with the aim of understanding option pricing, market behaviour, and investor risk aversion. These insights form the basis for many investment and risk management decisions. However, translating these insights from theory into practice still provides challenges. This dissertation distinguishes two categories of challenges: technical anomalies and market anomalies. The technical anomalies, arising from the inherent practical limitations of option trading, comprises discrete and truncated strikes, which introduce structural errors into the elegant mathematics of modern financial economics. Market anomalies on the other hand arise from overoptimistic speculation, and lead to irrational shapes of the volatility surface where fundamental assumptions of asset pricing are violated.

In chapter 2, technical anomalies where described as discernible gaps in strike prices and the constraints of minimum and maximum strikes. To evaluate

their impact on the analysis of the markets' perception of future variance, we derive an the expected value of future realized variance explicitly for a selection of common option pricing models. Using multiple sets of parameters for each model to simulate a wide range of market conditions and volatility surface shapes, we assess the accuracy of variance swap pricing methods against artificially discretized and truncated volatility surfaces. We find that discretization tends to be non-structural and interpolation along the strike grid leads to robust results. Truncation on the other hand leads to structural underestimation of future realized variance, and the commonly used extrapolation approaches introduce other sources of potentially large errors.

Chapter 3 examines the issue of truncation in the strike domain further. As the width of the strike grid fluctuates over time, the extend of the structural underestimation fluctuates in the same fashion. By utilizing a well-established result by Lee [2004] on the asymptotic slope of implied volatility at extreme strikes, we construct an explicit truncation compensation term that only depends on the strike and volatility of the outermost option contract for each wing. This compensation term measures the amount of missing future realized variance directly, and can be added to volatility indices such as the VIX Index with only minor modifications. Since investors are able to afford to purchase protection closer to the money, low volatility regimes induce larger truncation errors. During market corrections with peaking volatility the truncation error becomes negligible. A direct consequence of this is that strategies which rely on implied volatility for risk management are structurally over-leveraged.

Detecting irrational shapes of the volatility surface and their relation to speculative bubbles is the topic of chapter 4. When markets become irrational, core assumptions of financial theory fail to hold, and associated results are not applicable in these conditions. By explicitly specifying conditions where the fundamental theorem of asset pricing is violated within the framework of the SABR model, speculative bubbles can be identified. A large-scale study of 2576 stocks over 26 years reveals that the options market absorbs specu-

lative bubbles within a few days, while stock price bubbles have a lifetime in months or years. Based on events identified through martingale defects in the SABR model, we find that the distribution of asset returns changes, the number of actively traded option contracts as well as short interest increases, and institutional ownership decreases. These effects to be permanent and persistent. Our results further indicate that retail investors exhibit a higher propensity to participate in asset price bubbles than institutional investors. Over time, speculative bubbles appear almost exclusively in periods of positive price developments. Optimism fades quickly when markets correct.

In essence, this dissertation has unveiled the volatility surface as a dynamic and multifaceted construct that serves as a bridge between theoretical financial models and the complex practical realities of financial markets. It has demonstrated that the anomalies of the volatility surface are not mere curiosities but rather windows into the intricate dynamics of risk perception, market sentiment, and speculative behavior. These anomalies provide a rich landscape for future research, promising further advancements in the realms of option pricing, risk management, and the broader understanding of financial markets. As financial economics continues to evolve, the volatility surface remains a fertile ground for exploration and a testament to the enduring appeal of financial theory in practice.

Bibliography

Leif Andersen and Jesper Andreasen. Jump-Diffusion Processes: Volatility Smile Fitting and Numerical Methods for Option Pricing. *Diffusion Processes*, page 32, 2000.

Gurdip Bakshi, Charles Cao, and Zhiwu Chen. Pricing and hedging long-term options. *Journal of Econometrics*, 94(1-2):277–318, 2000. ISSN 03044076. doi: 10.1016/S0304-4076(99)00023-8.

Gurdip Bakshi, John Crosby, Xiaohui Gao, and Jinming Xue. Volatility Uncertainty, Disasters, and VIX Futures Contango. *SMU Cox School of Business Research Paper*, 21–19:46, 2021.

Gurdip Bakshi, John Crosby, and Xiaohui Gao. Dark Matter in (Volatility and) Equity Option Risk Premiums. *Operations Research*, 70(6):3108–3124, 2022.

Chris Bardgett, Elise Gourier, and Markus Leippold. Inferring volatility dynamics and risk premia from the S&P 500 and VIX markets. *Journal of Financial Economics*, 131(3):593–618, 2018. doi: 10.1016/j.jfineco.2018.09.008.

David S. Bates. Jumps and Stochastic Volatility: Exchange Rate Processes Implicit in Deutsche Mark Options. *The Review of Financial Studies*, 9 (1):69–107, 1996.

David S. Bates. Post-'87 crash fears in the S&P 500 futures option market. *Journal of Econometrics*, 94(1-2):181–238, 2000. ISSN 03044076. doi: 10.1016/S0304-4076(99)00021-4.

Shalom Benaim and Peter Friz. Regular Variation and Smile Asymptotics. *Mathematical Finance*, 19(1):1–12, 2009.

Shalom Benaim, Matthew Dodgson, and Dherminder Kainth. An arbitrage-free method for smile extrapolation. Technical report, Royal Bank of Scotland, 2009.

Shalom Benaim, Peter Friz, and Roger Lee. On Black-Scholes Implied Volatility at Extreme Strikes. In Rama Cont, editor, *Frontiers in Quantitative Finance*, pages 19–45. John Wiley & Sons, Inc., Hoboken, NJ, USA, January 2012. ISBN 978-1-118-26691-5 978-0-470-29292-1. doi: 10.1002/9781118266915.ch2.

Francesca Biagini, Hans Föllmer, and Sorin Nedelcu. Shifting martingale measures and the birth of a bubble as a submartingale. *Finance and Stochastics*, 18(2):297–326, April 2014. ISSN 0949-2984, 1432-1122. doi: 10.1007/s00780-013-0221-8.

Francesca Biagini, Andrea Mazzon, and Ari-Pekka Perkkiö. Optional projection under equivalent local martingale measures, March 2020.

Francesca Biagini, Lukas Gonon, Andrea Mazzon, and Thilo Meyer-Brandis. Detecting asset price bubbles using deep learning, October 2022.

Fischer Black and Myron Scholes. The Pricing of Options and Corporate Liabilities. 81(3):637–654, 1973. ISSN 0022-3808, 1537-534X. doi: 10.1086/260062. URL `https://www.journals.uchicago.edu/doi/10.1086/260062`.

Jesse Blocher, Xi Dong, Matthew C. Ringgenberg, and Pavel G. Savor. Short

Covering, November 2021.

Douglas T. Breeden and Robert H. Litzenberger. Prices of State-Contingent Claims Implicit in Option Prices. *The Journal of Business*, 51(4):621–651, 1978.

Mark Britten-Jones and Anthony Neuberger. Option Prices, Implied Price Processes, and Stochastic Volatility. *The Journal of Finance*, 55(2):839–866, 2000. ISSN 00221082. doi: 10.1111/0022-1082.00228.

Mark Broadie and Ashish Jain. Pricing and Hedging Volatility Derivatives. *The Journal of Derivatives*, 15(3):7–24, 2008a. ISSN 1074-1240, 2168-8524. doi: 10.3905/jod.2008.702503.

Mark Broadie and Ashish Jain. The Effect of Jumps and Discrete Sampling on Volatility and Variance Swaps. *International Journal of Theoretical and Applied Finance*, 11(8):761–797, 2008b. ISSN 0219-0249, 1793-6322. doi: 10.1142/S0219024908005032.

Colin Camerer. Bubbles and Fads in Asset Prices. *Journal of Economic Surveys*, 3(1):3–41, March 1989. ISSN 0950-0804, 1467-6419. doi: 10.1111/j.1467-6419.1989.tb00056.x.

Peter Carr and Dilip B. Madan. Towards a Theory of Volatility Trading. In Robert A. Jarrow, editor, *Volatility: New Estimation Techniques for Pricing Derivatives*, pages 417–427. RISK Books, London, 1998.

Peter Carr and Liuren Wu. A Tale of Two Indices. *The Journal of Derivatives*, (2):13–29, 2006.

Peter Carr and Liuren Wu. Variance Risk Premiums. *Review of Financial Studies*, 22(3):1311–1341, March 2009. ISSN 0893-9454, 1465-7368. doi: 10.1093/rfs/hhn038.

Peter Carr and Liuren Wu. Analyzing volatility risk and risk premium in

option contracts: A new theory. *Journal of Financial Economics*, 120(1): 1–20, 2016. ISSN 0304405X. doi: 10.1016/j.jfineco.2016.01.004.

Peter Carr, Roger Lee, and Liuren Wu. Variance swaps on time-changed Lévy processes. *Finance and Stochastics*, 16(2):335–355, April 2012. ISSN 0949-2984, 1432-1122. doi: 10.1007/s00780-011-0157-9.

CBOE. CBOE to Apply VIX Methodology to Individual Equity Options. https://ir.cboe.com/news-and-events/2011/01-05-2011/cboe-apply-vix-methodology-individual-equity-options, 2011.

CBOE. Double the Fun with CBOE's VVIX Index. https://cdn.cboe.com/resources/indices/documents/vvix-termstructure.pdf, 2012.

CBOE. Cboe VIX Whitepaper - Cboe Volatility Index. https://cdn.cboe.com/resources/futures/vixwhite.pdf, 2018.

Hui Chen, Winston Wei Dou, and Leonid Kogan. Measuring 'Dark Matter' in Asset Pricing Models, January 2022.

Ing-Haw Cheng. The VIX Premium. *The Review of Financial Studies*, 32 (1):180–227, January 2019. ISSN 0893-9454, 1465-7368. doi: 10.1093/rfs/hhy062.

Patrick Cheridito, Damir Filipovic, and Robert L. Kimmel. Market Price of Risk Specifications for Affine Models: Theory and Evidence. *Journal of Financial Economics*, 83:123–170, 2007. ISSN 1556-5068. doi: 10.2139/ssrn.524724.

Jaehyuk Choi, Kwangmoon Kim, and Minsuk Kwak. Numerical approximation of the implied volatility under arithmetic Brownian motion. *Applied Mathematical Finance*, 16(3), 2009.

John R. Conlon. Simple Finite Horizon Bubbles Robust to Higher Order

Knowledge. *Econometrica*, 72(3):927–936, May 2004. ISSN 0012-9682, 1468-0262. doi: 10.1111/j.1468-0262.2004.00516.x.

Rama Cont and José da Fonseca. Dynamics of Implied Volatility Surfaces. *Quantitative Finance*, 2:45–60, 2002.

Alexander M. G. Cox and David G. Hobson. Local martingales, bubbles and option prices. *Finance and Stochastics*, 9(4):477–492, October 2005. ISSN 0949-2984, 1432-1122. doi: 10.1007/s00780-005-0162-y.

John C. Cox, Stephen A. Ross, and Mark Rubinstein. Option pricing: A simplified approach. *Journal of Financial Economics*, 7(3):229–263, September 1979. ISSN 0304405X. doi: 10.1016/0304-405X(79)90015-1.

Freddy Delbaen and Walter Schachermayer. A general version of the fundamental theorem of asset pricing. *Mathematische Annalen*, 300(1):463–520, September 1994. ISSN 0025-5831, 1432-1807. doi: 10.1007/BF01450498.

Kresimir Demeterfi, Emanuel Derman, Michael Kamal, and Joseph Zou. More than you ever wanted to know about volatility swaps. *Goldman Sachs Quantitative Strategies Research Notes*, 1999.

Emanuel Derman and Michael B. Miller. *The Volatility Smile: An Introduction for Students and Practitioners*. The Wiley Finance Series. Wiley, Hoboken, New Jersey, 2016. ISBN 978-1-118-95918-3 978-1-118-95917-6.

Maneesh S. Deshpande, Arnab Sen, Brian Pu, Si Gao, and Elias Krauklis. Impact of Retail Options Trading. Technical report, Barclays US Equity Derivatives Strategy, September 2020.

Adrian A Drăgulescu and Victor M Yakovenko. Probability distribution of returns in the Heston model with stochastic volatility. *Quantitative Finance*, 2(6):443–453, 2002. ISSN 1469-7688, 1469-7696. doi: 10.1080/14697688.2002.0000011.

Bruno Dupire. Pricing with a Smile. *RISK*, 7, 1994.

Bjorn Eraker. The Volatility Premium. *The Quarterly Journal of Finance*, 11(03):2150014, September 2021. ISSN 2010-1392, 2010-1406. doi: 10.1142/S2010139221500142.

Bjorn Eraker and Aoxiang Yang. The Price of Higher Order Catastrophe Insurance: The Case of VIX Options. *SSRN Electronic Journal*, 2020. ISSN 1556-5068. doi: 10.2139/ssrn.3520256.

Muslim Farooque. 7 Tech Stocks to Buy That Could 10X by 2025, September 2022.

Matthias R. Fengler. Arbitrage-free smoothing of the implied volatility surface. *Quantitative Finance*, 9(4):417–428, 2009. ISSN 1469-7688, 1469-7696. doi: 10.1080/14697680802595585.

Matthias R. Fengler and Lin-Yee Hin. Semi-Nonparametric Estimation of the Call Price Surface Under No-Arbitrage Constraints. *Journal of Econometrics*, 184:242–261, 2015.

Marcelo Fernandes, Marcelo C. Medeiros, and Marcel Scharth. Modeling and predicting the CBOE market volatility index. *Journal of Banking & Finance*, 40:1–10, March 2014. ISSN 03784266. doi: 10.1016/j.jbankfin.2013.11.004.

Stephen Figlewski. Estimating the Implied Risk Neutral Density for the U.S. Market Portfolio. In Tim Bollerslev, Jeffrey Russell, and Mark Watson, editors, *Volatility and Time Series Econometrics: Essays in Honor of Robert F. Engle*. Oxford University Press, 2008. ISBN 978-0-19-954949-8. doi: 10.1093/acprof:oso/9780199549498.001.0001.

Jeff Fleming, Barbara Ostdiek, and Robert E. Whaley. Predicting stock market volatility: A new measure. *Journal of Futures Markets*, 15(3):265–

302, May 1995. ISSN 02707314, 10969934. doi: 10.1002/fut.3990150303.

Danielle Florens-Zmirou. On estimating the diffusion coefficient from discrete observations. *Journal of Applied Probability*, 30(4):790–804, December 1993. ISSN 0021-9002, 1475-6072. doi: 10.2307/3214513.

M. Fukasawa, I. Ishida, N. Maghrebi, K. Oya, M. Ubukata, and K. Yamazaki. Model-free implied volatility: From Surface to Index. *International Journal of Theoretical and Applied Finance*, 14(04):433–463, 2011. ISSN 0219-0249, 1793-6322. doi: 10.1142/S0219024911006681.

Nicola Fusari, Robert Jarrow, and Sujan Lamichhane. Testing for Asset Price Bubbles using Options Data, 2022.

Jim Gatheral. *The Volatility Surface: A Practitioner's Guide*. Wiley Finance Series. John Wiley & Sons, Hoboken, N.J, 2006. ISBN 978-0-471-79251-2.

Kathrin Glau, Paul Herold, Dilip B. Madan, and Christian Pötz. The Chebyshev method for the implied volatility. *arXiv:1710.01797 [q-fin]*, 2017.

Peter H. Gruber, Claudio Tebaldi, and Fabio Trojani. The Price of the Smile and Variance Risk Premia. *Management Science*, 67(7):4056–4074, 2020. ISSN 0025-1909, 1526-5501. doi: 10.1287/mnsc.2020.3689.

Archil Gulisashvili. Asymptotic Formulas with Error Estimates for Call Pricing Functions and the Implied Volatility at Extreme Strikes. *SIAM Journal on Financial Mathematics*, 1(1):609–641, January 2010. ISSN 1945-497X. doi: 10.1137/090762713.

Patrick S. Hagan, Deep Kumar, Andrew Lesniewski, and Diana Woodward. Managing Smile Risk. *Willmott Magazine*, pages 84–108, 2002.

Steven Heston. A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options. *The Review of Financial Studies*, 6(2):327–343, 1993.

Steven L. Heston, Mark Loewenstein, and Gregory A. Willard. Options and Bubbles. *Review of Financial Studies*, 20(2):359–390, March 2007. ISSN 0893-9454, 1465-7368. doi: 10.1093/rfs/hhl005.

Hardy M. Hodges. Arbitrage Bounds of the Implied Volatility Strike and Term Structures of European-Style Options. *The Journal of Derivatives*, 3(4):23–35, 1996. ISSN 1074-1240, 2168-8524. doi: 10.3905/jod.1996.407950.

Robert Jarrow and Simon Kwok. Inferring Financial Bubbles from Option Data, June 2021.

Robert Jarrow and Philip Protter. Positive Alphas, Abnormal Performance, and Illusory Arbitrage. *Mathematical Finance*, 23(1):39–56, January 2013. ISSN 09601627. doi: 10.1111/j.1467-9965.2011.00489.x.

Robert A. Jarrow. Market Manipulation, Bubbles, Corners, and Short Squeezes. *The Journal of Financial and Quantitative Analysis*, 27(3):311, September 1992. ISSN 00221090. doi: 10.2307/2331322.

Robert A. Jarrow. Asset Price Bubbles, December 2015.

Robert A. Jarrow and Philip Protter. The Martingale Theory of Bubbles: Implications for the Valuation of Derivatives and Detecting Bubbles, May 2010.

Robert A. Jarrow, Philip Protter, and Kazuhiro Shimbo. Asset Price Bubbles in Complete Markets. *Advances in Mathematical Finance*, pages 97–121, 2007. doi: 10.1007/978-0-8176-4545-8_7.

Robert A. Jarrow, Philip Protter, and Kazuhiro Shimbo. Asset Price Bubbles in Incomplete Markets. *Mathematical Finance*, 20(2):145–185, April 2010. ISSN 09601627, 14679965. doi: 10.1111/j.1467-9965.2010.00394.x.

Robert A. Jarrow, Younes Kchia, and Philip Protter. How to Detect an Asset

Bubble, February 2011.

George J. Jiang and Yisong S. Tian. The Model-Free Implied Volatility and Its Information Content. *Review of Financial Studies*, 18(4):1305–1342, 2005. ISSN 0893-9454, 1465-7368. doi: 10.1093/rfs/hhi027.

George J. Jiang and Yisong S. Tian. Extracting Model-Free Volatility from Option Prices: An Examination of the Vix Index. *Journal of Derivatives*, 14(3), 2007. ISSN 1556-5068. doi: 10.2139/ssrn.880459.

Nabil Kahale. Smile interpolation and calibration of the local volatility model. *Working Paper*, 2005.

Márcio Poletti Laurini. Imposing no-arbitrage conditions in implied volatilities using constrained smoothing splines. *Applied Stochastic Models in Business and Industry*, 27(6):649–659, 2011. ISSN 15241904. doi: 10.1002/asmb.877.

Fabien Le Floc'h and Gary J. Kennedy. Explicit SABR Calibration Through Simple Expansions, July 2014.

Roger Lee. The Moment Formula for Implied Volatility at Extreme Strikes. *Mathematical Finance*, 14(3):469–480, 2004. ISSN 0960-1627, 1467-9965. doi: 10.1111/j.0960-1627.2004.00200.x.

Roger Lee. Implied Volatility: Statics, Dynamics, and Probabilistic Interpretation. In Ricardo Baeza-Yates, Joseph Glaz, Henryk Gzyl, Jürgen Hüsler, and José Luis Palacios, editors, *Recent Advances in Applied Probability*, pages 241–268. Kluwer Academic Publishers, Boston, 2005. ISBN 978-0-387-23378-9.

Mark Loewenstein and Gregory A. Willard. Local martingales, arbitrage, and viability. *Economic Theory*, 16(1):135–161, July 2000. ISSN 0938-2259, 1432-0479. doi: 10.1007/s001990050330.

Richard B. Loucks. Chebyshev Polynomial Fit for Terrain Elevation. Technical report, Defense Technical Information Center, Fort Belvoir, VA, 1996.

Robert Merton. Option Pricing when Underlying Stock Prices are Discontinuous. *Journal of Financial Economics*, 3:125–144, 1976.

Robert C. Merton. Theory of rational option pricing. *The Bell Journal of Economics and Management Science*, 4(1):141–183, 1973. ISSN 00058556. URL `http://www.jstor.org/stable/3003143`.

Hannes Mohrschladt and Judith C. Schneider. Idiosyncratic volatility, option-based measures of informed trading, and investor attention. *Review of Derivatives Research*, 24(3):197–220, October 2021. ISSN 1573-7144. doi: 10.1007/s11147-021-09175-7.

MSCI. MSCI USA IMI Sector Indexes. Technical report, MSCI, 2017.

J. A. Nelder and R. Mead. A Simplex Method for Function Minimization. *The Computer Journal*, 7(4):308–313, January 1965. ISSN 0010-4620, 1460-2067. doi: 10.1093/comjnl/7.4.308.

Anthony Neuberger. The Log Contract. *Journal of Portfolio Management*, 20(2):74–80, 1994.

Yoshiki Obayashi, Philip Protter, and Shihao Yang. The Lifetime of a Financial Bubble, March 2016.

Petteri Piiroinen, Lassi Roininen, Tobias Schoden, and Martin Simon. Asset Price Bubbles: An Option-based Indicator, July 2018.

Petteri Piiroinen, Lassi Roininen, and Martin Simon. Brexit Risk Implied by the SABR Martingale Defect in the EUR-GBP Smile. *Working Paper*, 2019.

Philip Protter. Mathematical models of bubbles. *Quantitative Finance*

Letters, 4(1):10–13, January 2016. ISSN 2164-9502, 2164-9510. doi: 10.1080/21649502.2015.1165863.

Sidney Resnik. *A Probability Path*. Birkhauser Boston, New York, 1998.

Stephen A. Ross. Finance. In Palgrave Macmillan, editor, *The New Palgrave Dictionary of Economics*. 1987.

SEC. Staff Report on Equity and Options Market Structure Conditions in Early 2021. Technical report, Securities Exchange Commision, 2021.

Aaron Smith. Experimental drug could be first to stall Alzheimers. https://money.cnn.com/2007/08/06/news/companies/myriad/index.htm, August 2007.

Zhiyuan Sun. Is Sundial Growers a Buy? https://www.fool.com/investing/2021/05/14/is-sundial-growers-a-buy/, May 2021.

Appendix A

Chapter 2

A.1 Calculation of the VIX Index

CBOE [2018] lays out the methodology of the VIX Index in detail. It fundamentally consists of the following steps:

1. The two maturities closest to the 30 days-to-expiration mark are selected and all relevant options are chosen.

2. Of those options, options with zero-bids are removed; If two options with adjacent strikes have zero-bids, all options further from the money are removed as well.

3. The MFIV is computed for both maturities.

4. The two resulting volatility estimates are linearly interpolated to match the desired time-to-expiration of 30 days.

A.2 MFIV in terms of log-forward moneyness

We begin with the definition by Britten-Jones and Neuberger [2000], equation 3.2 as

$$\mathbb{E}_0^{\mathbb{Q}}\left[\int_0^T \left(\frac{dF_t}{F_t}\right)^2\right] = 2\int_0^\infty \frac{C(K,\tau,\sigma_K) - max(F_0 - K, 0)}{K^2} dK. \tag{A.1}$$

By definition, we have

$$k = log\left(\frac{K}{F_0}\right) \tag{A.2}$$

$$\Leftrightarrow \quad K = F_0 e^k \tag{A.3}$$

and thus $\sigma_K = \sigma_k$ to shorten notation.

The price of an option in log-forward moneyness is quoted in relative terms, thus

$$C(K,\tau,\sigma_K) = F_0 c(k,\tau,\sigma_k). \tag{A.4}$$

Substitute and simplify to

$$MFIV = 2\int_{-\infty}^\infty \frac{F_0 c(k,\tau,\sigma_k) - max(F_0 - F_0 e^k, 0)}{(F_0 e^k)^2} F_0 e^k dk \tag{A.5}$$

$$= 2\int_{-\infty}^\infty \frac{c(k,\tau,\sigma_k) - max(1 - e^k, 0)}{e^k} dk. \tag{A.6}$$

A.3 Derivation of the extrapolated Black-Scholes call price

In the standard Black-Scholes model, the price of a call option is given by

$$c(k,\tau,\sigma) = \Psi(d_1(k,\tau,\sigma)) - e^k\Psi(d_2(k,\tau,\sigma)), \tag{A.7}$$

$$d_1(k,\tau,\sigma) = -\frac{k}{\sigma\sqrt{\tau}} + \tfrac{1}{2}\sigma\sqrt{\tau}, \text{ and} \tag{A.8}$$

$$d_2(k,\tau,\sigma) = d_1(k,\tau,\sigma) - \sigma\sqrt{\tau}, \tag{A.9}$$

where Ψ is the normal cumulative distribution function (CDF).

By Lee [2004], as $k \longrightarrow \pm\infty$, $\sigma_{BSIV} \longrightarrow \sqrt{\frac{\beta}{\tau}|k|}$, where $\beta \in [0,2]$ being specific to either the left or right wing. To extrapolate c in the tails, compute β from the outermost option price, then substitute σ to get the price of a far-out-of-the-money call option in line with no-arbitrage limits as

$$\tilde{c}(k,\tau,\beta) = c(k,\tau,\sqrt{\frac{\beta}{\tau}|k|}) \tag{A.10}$$

$$= \Psi(\tilde{d}_1(k,\tau)) - e^k\Psi(\tilde{d}_2(k,\tau)), \tag{A.11}$$

$$\tilde{d}_1(k,\tau,\beta) = d_1(k,\tau,\sqrt{\frac{\beta}{\tau}|k|}) \tag{A.12}$$

$$= -\frac{k}{\sqrt{\beta|k|}} + \tfrac{1}{2}\sqrt{\beta|k|}, \text{ and} \tag{A.13}$$

$$\tilde{d}_2(k,\tau,\beta) = d_2(k,\tau,\sqrt{\frac{\beta}{\tau}|k|}) \tag{A.14}$$

$$= \tilde{d}_1(k,\tau,\beta) - \sqrt{\beta|k|}. \tag{A.15}$$

Repeat for the price of a put option.

A.4 Dynamics of Lee [2004]'s tail parameter

Figure A.1 shows the β-curves for six different models to illustrate the general stability of the tail parameter dynamics. With short time-to-expiration, the tail parameters level out quickly. For longer time-to-expiration the specific model properties and parameters become more important.

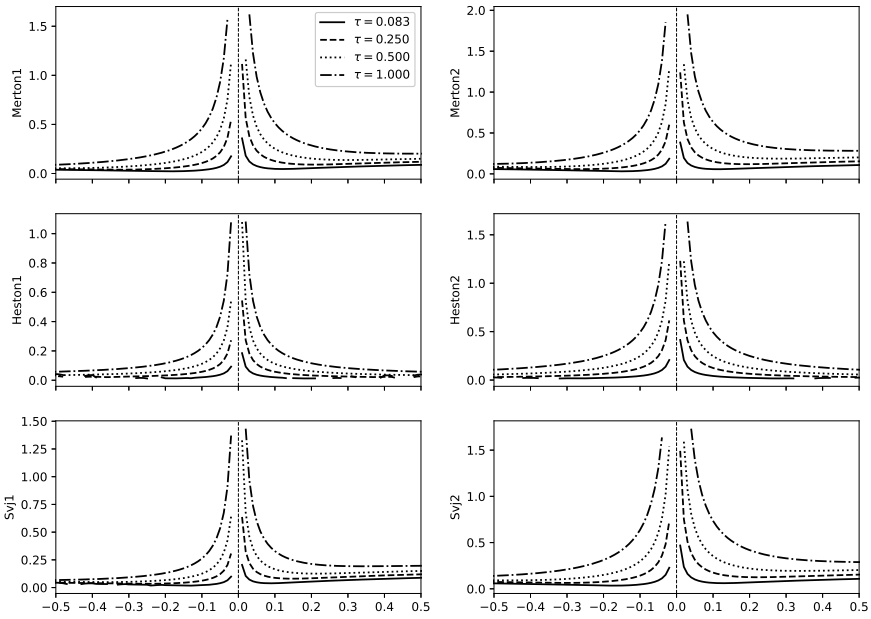

Figure A.1: β-curves in comparison. The top row shows the results for the Merton-model with $\Omega_{Merton1} = \{\sigma = 0.2, \alpha = 0.5, \delta = 0.5, \lambda = 0.02\}$ and $\Omega_{Merton2} = \{\sigma = 0.2, \alpha = 0.3, \delta = 0.7, \lambda = 0.04\}$. The middle row shows the results for the Heston-model with $\Omega_{Heston1} = \{v_0 = 0.0225, \bar{v} = 0.0225, \kappa = 3, \eta = 0.25, \rho = 0\}$ and $\Omega_{Heston2} = \{v_0 = 0.09, \bar{v} = 0.09, \kappa = 3, \eta = 0.25, \rho = 0\}$. The bottom row shows the results for the Stochastic Volatility with Jumps (SVJ)-model with $\Omega_{Svj1} = \{\Omega_{Merton1}, \Omega_{Heston1}\}$ and $\Omega_{Svj2} = \{\Omega_{Merton2}, \Omega_{Heston2}\}$.

A.5 Truncation error under different cutoff strikes

Figure A.2 provides the bottom plot of figure 3.3 for the minimum cutoff strikes ±0.05, ±0.075, and ±0.1. While there are some visible differences, the general result appears to hold.

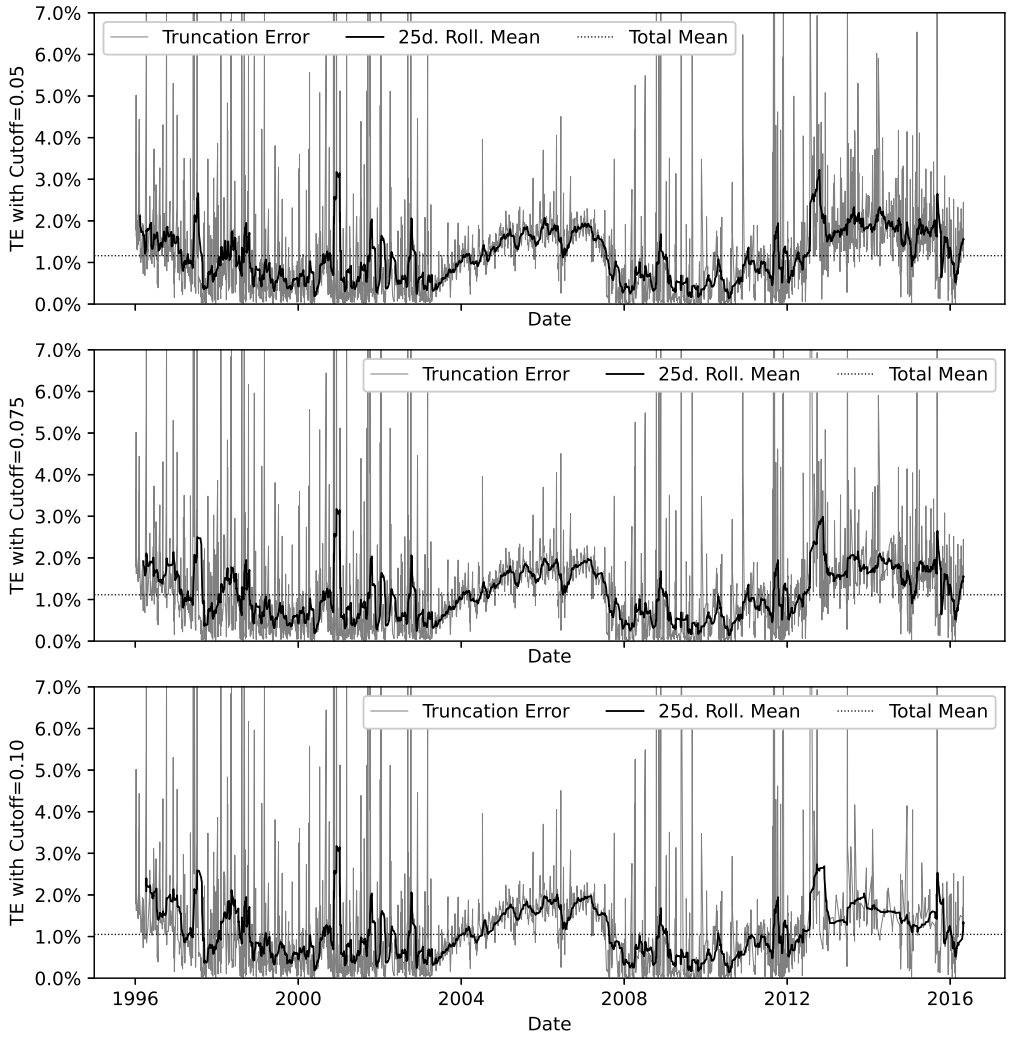

Figure A.2: Comparison of truncation error for the minimum cutoff strikes ±0.05, ±0.075, and ±0.1.

Appendix B

Chapter 3

B.1 Calibration results

Table B.1 provides descriptive statistics of calibration results. The large sample size leads to a small number of calibrations with above-desirable $RSME_{IV}$ figures, but for the vast majority of observations these results indicate adequate fit. Comparing calibration results for the full sample with those observations where the resulting $A(\cdot)$ exceeds a set of thresholds confirms that the calibration achieves acceptable fit most of the time even under martingale defect conditions.

B.2 Total trading activity over time

Within the timespan of our dataset, option trading activity has increased. The number of underlyings with actively traded options as well as the number of options per underlying have both grown. This development propagates into our results, since more underlyings with active option trading implies a higher number of possible events. To measure and compensate for this

		Full Sample	$A(\cdot) > 0$	$A(\cdot) \geq 0.01$	$A(\cdot) \geq 0.05$	$A(\cdot) \geq 0.1$
Count	total	1,402,230	159,309	109,414	56,722	39,270
	% of total		11.36%	7.80%	4.05%	2.80%
$RSME_{IV}$	mean	0.0097	0.0192	0.0182	0.0199	0.0219
	min	0.0000	0.0000	0.0000	0.0000	0.0000
	25% quantile	0.0000	0.0000	0.0000	0.0000	0.0017
	50% quantile	0.0027	0.0080	0.0072	0.0094	0.0116
	75% quantile	0.0102	0.0255	0.0238	0.0259	0.0280
	max	0.2121	0.2121	0.2121	0.2121	0.2121
Defect $A(\cdot)$	mean	0.0172	0.1511	0.2181	0.3978	0.5432
	min	0.0000	0.0000	0.0100	0.0500	0.1000
	25% quantile	0.0000	0.0071	0.0228	0.0857	0.1743
	50% quantile	0.0000	0.0245	0.0535	0.1886	0.4281
	75% quantile	0.0000	0.0973	0.2038	0.8419	1.0000
	max	1.0000	1.0000	1.0000	1.0000	1.0000

Table B.1: Descriptive statistics of calibration results. The first two rows provide the number of valid calibrations in the sample, as well as their proportions. The second set of rows provides statistics of calibration $RSME_{IV}$. The third set of rows provides statistics of the resulting martingale defect indicator $A(\cdot)$. The columns provide statistic conditional on wether $A(\cdot)$ exceeds a given threshold $[0.0, 0.01, 0.05, 0.1]$.

trend, we use the total number of actively traded option contracts as proxy for market-wide option trading activity. We employ the same liquidity requirements as laid out in section 4.4. Figure 1.1 provides the total number of daily active contracts for the SPX Index as well as the complete securities universe. As activity is fluctuating strongly from day to day, we use 252 trading day rolling averages. Figure B.1 provides the average number of contracts available for the calibration of the SABR model.

B.3 Relative trading activity over time by maturity

Appendix B.2 examines the total number of actively traded contracts. In this section, we will examine the proportional trading activity by maturity, as far as it is relevant to our analysis. We separate contracts by their remaining

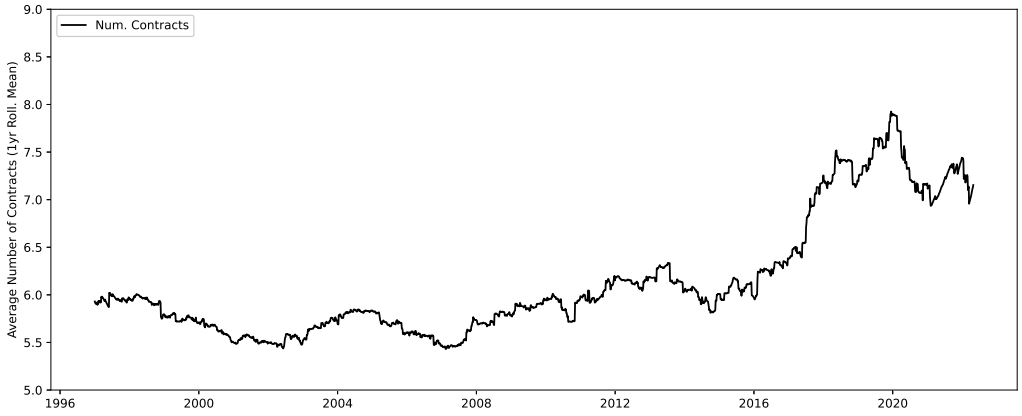

Figure B.1: Average daily number of contracts available for calibration. For visual clarity, the plot shows one year rolling averages.

lifetime in days. These slices are commonly called *tenors*. Table B.2 provides a list of tenor definitions.

Tenor	1 week	2 weeks	1 month	3 months	6 months
DTE	$2-7$ days	$12-16$ days	$25-35$ days	$80-100$ days	$165-195$ days

Table B.2: Tenor classifications.

Figure B.2 provides an overview of the development of available option contracts within a tenor. It shows the number of contracts within a range of days-to-maturity relative to the total number of contracts available. We apply the liquidity requirements laid out in section 4.4. As the periods are not exactly adjacent, the sum of the relative proportions may be less than 100%. Each day, all available option contracts are sorted by days to expiration, and their proportions are compared over time. The proportion of available options with a lifetime of less than one month begins to increase in 2012. The proportion of options with a lifetime of approximately one month remains very roughly constant, while the proportion of longer options declines from 2012 onward. After 2012, the proportion of available contracts has increased strongly. The proportion of the one-month tenor is relatively stable, but has increased as well. Overall, the market appears to have embraced the availability of low-DTE contracts. This fundamental shift in the dataset motivates us to provide our findings based on the one month tenor as well as

the full surface.

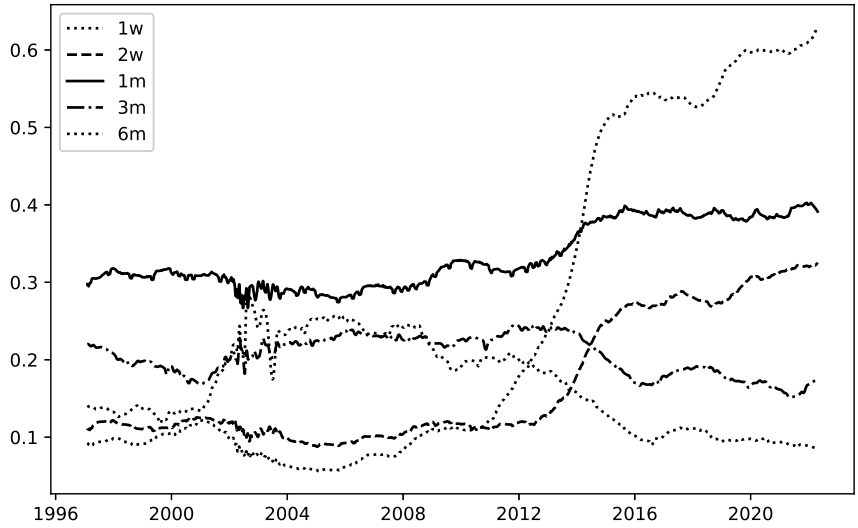

Figure B.2: Proportions of traded option contracts by tenor. For visual clarity, the plot shows one year rolling averages.

B.4 Persistence

Table B.3 provides details on the results discussed in section 4.4. Furthermore, the number of affected stocks is disclosed.

B.5 Results for Full Surface

Table B.4 complements table 4.2 and provides results based on the full surface.

Table B.5 complements table 4.3 and provides results based on the full sur-

by Threshold		by Persistence					
		1 Month Tenor			Full Surface		
		1 day	2 days	3 days	1 day	2 days	3 days
0.01	Events	18526	3505	1339	38798	6731	2477
	Companies	1591	820	475	1923	1193	717
0.05	Events	9983	1518	529	21087	2745	856
	Companies	1391	527	249	1806	806	382
0.10	Events	6659	823	274	14602	1519	390
	Companies	1203	353	149	1674	561	217

Table B.3: Number of events and number of unique affected companies for different persistence and threshold configurations. The left half provides results based on options with approximately one month to expiration, the right half provides results based on the entire volatility surface. The upper row provides the number of events in total, the lower row shows across how many companies these events are spread out. A persistence of one day means that the indicator must be above threshold for one day, i.e. there is no persistence requirement.

face.

Table B.6 complements table 4.4 and provides results based on the full surface.

Period t_{event}	Full Surface						One-month Tenor					
	N	t-statistics	p-values				N	t-statistics	p-values			
		Mean	$p \leq 5\%$		$p \leq 1\%$			Mean	$p \leq 5\%$		$p \leq 1\%$	
21	1149	0.19	22.55%	(7.13%)	12.19%	(3.17%)	780	0.38	10.01%	(2.16%)	4.58%	(0.86%)
63	1196	0.36	32.41%	(9.81%)	23.34%	(5.60%)	953	0.18	20.63%	(5.81%)	12.13%	(2.76%)
84	1195	0.45	33.86%	(17.50%)	24.66%	(12.73%)	930	0.32	19.19%	(9.85%)	10.50%	(6.67%)
105	1168	0.67	35.82%	(19.14%)	28.28%	(13.44%)	944	0.44	23.08%	(10.18%)	14.73%	(6.19%)
126	1179	0.79	39.96%	(22.60%)	31.25%	(16.67%)	956	0.49	25.74%	(14.27%)	17.34%	(9.92%)
189	1152	0.95	42.70%	(23.72%)	34.07%	(17.96%)	957	0.61	28.42%	(14.42%)	20.40%	(9.56%)
252	1090	1.35	45.45%	(29.85%)	37.25%	(24.20%)	912	0.68	31.13%	(20.06%)	22.04%	(13.84%)

Table B.4: Aggregate results for t-tests of the number of actively traded option contracts before and after a martingale defect event, based on the calibration of the SABR model against the full volatility surface. Event Period t_{event} provides the number of trading days before and after an event under consideration. Results under *Full Surface* count all available options, while results under *One-month Tenor* count only contracts within the one-month tenor. N refers to the number of events after discarding stocks with insufficient data and overlapping event periods. Mean t-statistic reports the mean of the t-statistic of all events. $p \leq 5\%$ and $p \leq 1\%$ report the quantile of p-values in the sample. The corresponding results of the placebo study with shuffled event dates are shown in parentheses.

Period t_{event}	N	t-statistics	p-values			
		Mean	$p \leq 5\%$		$p \leq 1\%$	
21	1060	2.185	75.2%	(46.5%)	66.7%	(38.5%)
63	1042	1.250	77.9%	(64.4%)	71.8%	(64.2%)
84	1033	0.979	80.8%	(73.7%)	75.4%	(63.1%)
105	1010	0.867	81.0%	(60.9%)	75.7%	(48.1%)
126	985	0.647	83.1%	(79.5%)	77.9%	(66.5%)
189	950	0.879	86.3%	(73.8%)	81.5%	(63.5%)
252	884	0.998	88.0%	(75.1%)	83.7%	(63.2%)

Table B.5: Aggregate results for t-tests of outstanding short interest divided by ADTV before and after a martingale defect event, based on the calibration of the SABR model against the full volatility surface. Event Period t_{event} provides the number of trading days before and after an event under consideration. N refers to the number of events after discarding stocks with insufficient data and overlapping event periods. Mean t-statistic reports the mean of the t-statistic of all events. $p \leq 5\%$ and $p \leq 1\%$ report the quantile of p-values in the sample. The corresponding results of the placebo study with shuffled event dates are shown in parentheses.

Period t_{event}	N	t-statistics	p-values			
		Mean	$p \leq 5\%$		$p \leq 1\%$	
21	1060	-0.264	59.0%	(43.1%)	52.9%	(34.2%)
63	1042	-0.479	65.1%	(49.1%)	62.9%	(43.8%)
84	1033	-0.176	64.5%	(48.6%)	62.0%	(49.4%)
105	1010	-0.319	63.8%	(57.4%)	61.4%	(54.0%)
126	985	-0.244	64.0%	(58.5%)	61.7%	(56.4%)
189	950	-0.308	65.4%	(57.4%)	63.3%	(53.5%)
252	884	-0.102	64.7%	(59.6%)	63.1%	(56.8%)

Table B.6: Aggregate results for t-tests of the percentage of institutional ownership before and after a martingale defect event, based on the calibration of the SABR model against the full volatility surface. Event Period t_{event} provides the number of trading days before and after an event under consideration. N refers to the number of events after discarding stocks with insufficient data and overlapping event periods. Mean t-statistic reports the mean of the t-statistic of all events. $p \leq 5\%$ and $p \leq 1\%$ report the quantile of p-values in the sample. The corresponding results of the placebo study with shuffled event dates are shown in parentheses.

Declaration

Statement pursuant to § 22 paragraph 7 and § 23 paragraph 7 of APB TU Darmstadt

I herewith formally declare that I, Philip Stahl, have written the submitted dissertation independently pursuant to § 22 paragraph 7 of APB TU Darmstadt. I did not use any outside support except for the quoted literature and other sources mentioned in the dissertation. I clearly marked and separately listed all of the literature and all of the other sources which I employed when producing this academic work, either literally or in content. This dissertation has not been submitted for any other degree or professional qualification.

I am aware, that in case of an attempt at deception based on plagiarism (§38 Abs. 2 APB), the dissertation would be counted as one failed examination attempt.

Frankfurt am Main, 11.12.2023

Philip Stahl