Universität Hohenheim
Institut für Volkswirtschaftslehre
Fachgebiet Statistik und Ökonometrie I

Forecasting DAX Volatility: A Comparison of Time Series Models and Implied Volatilities

Dissertation
zur Erlangung des Grades eines
Doktors der Wirtschaftswissenschaften
(Dr. oec.)

vorgelegt der
Fakultät Wirtschafts- und Sozialwissenschaften
der Universität Hohenheim

von

Dipl. oec. Harald Weiß
Stuttgart-Hohenheim
2016

Diese Arbeit wurde im Jahr 2016 von der Fakultät Wirtschafts- und Sozialwissenschaften der Universität Hohenheim als Dissertation zur Erlangung des Grades eines Doktors der Wirtschaftswissenschaften (Dr. oec.) angenommen.

Datum der mündlichen Doktorprüfung: 12. September 2016

Dekan: Prof. Dr. Dirk Hachmeister
Prüfungsvorsitz: Prof. Dr. Aderonke Osikominu
Erstgutachter: Prof. Dr. Gerhard Wagenhals
Zweitgutachter: Prof. Dr. Robert Jung

Vorwort

Die vorliegende Arbeit wurde im Jahr 2016 von der Fakultät Wirtschafts- und Sozialwissenschaften der Universität Hohenheim als Dissertation angenommen.

Mein besonderer Dank gilt meinem Doktorvater Prof. Dr. Gerhard Wagenhals, der mir zunächst als externer Doktorand und später als Mitarbeiter die Promotion ermöglicht hat. Herrn Prof. Dr. Robert Jung danke ich für seine kritischen Anmerkungen und das Erstellen des Zweitgutachtens. Auch Frau Prof. Dr. Aderonke Osikominu gilt ein herzliches Dankeschön für die Mitwirkung am Promotionsverfahren.

Bei meinen Kolleginnen und Kollegen vom Lehrstuhl für Statistik und Ökonometrie möchte ich mich nicht nur für die fachliche Unterstützung, sondern vor allem auch für das sehr angenehme Arbeitsklima bedanken. Hier sind Prof. Dr. Ulrich Scheurle, Martina Rabe, Dr. Frauke Wolf, Dr. Steffen Wirth, Dr. Katja Hölsch, Dr. Sebastian Moll, Dr. Wolf-Dieter Heinbach, Dr. Ulrike Berberich, Dr. Stephanie Schröpfer und Prof. Dr. Gröner zu nennen. Ergänzend dazu bin ich Dr. Robert Maderitsch zu Dank verfplichtet, mit dem ich zahlreiche Gespräche über Realisierte Volatilitäten, Strukturbrüche und vieles mehr geführt habe. Ebenso bedanke ich mich bei unseren ehemaligen wissenschaftlichen Hilfskräften Claudia Illgen, Marc Epple und Felix Prettl. Sie haben mich bei der Erstellung der Grafiken und beim Korrekturlesen tatkräftig unterstützt.

Mein größter Dank gilt meiner Frau Simone Weiß und meiner Familie. Eure liebevolle Unterstützung und Euer Rückhalt waren mir eine wichtige Stütze während der Promotion. Die vorliegende Arbeit widme ich meinen Eltern, die mein Studium und die Promotion nach besten Kräften unterstützt haben.

Stuttgart, Dezember 2016 Harald Weiß

Contents

List of Tables

List of Figures

List of Abbreviations

AC	autocorrelation
ACF	autocorrelation function
ADF	augmented Dickey-Fuller
AGARCH	Asymmetric GARCH
AIC	Akaike information criterion
AJD	affine jump-diffusion
AR	Autoregressive
ARCH	Autoregressive Conditional Heteroskedasticity
ARCH LM	Autoregressive Conditional Heteroskedasticity Lagrange multiplier
ARFIMA	Autoregressive Fractionally Integrated Moving Average
ARIMA	Autoregressive Integrated Moving Average
ARMA	Autoregressive Moving Average
ATM	at-the-money
BS	Black-Scholes
BS PDE	Black-Scholes partial differential equation
CBOE	Chicago Board Options Exchange
CEV	constant elasticity of variance
CME	Chicago Mercantile Exchange
DAX	Deutscher Aktienindex
DFGLS	Dickey-Fuller generalised least-squares
DJIA	Dow Jones Industrial Average

DM	Diebold-Mariano
DTB	Deutsche Terminbörse
EGARCH	Exponential GARCH
EIV	errors-in-variables
EONIA	Euro OverNight Index Average
EPA	equal predictive ability
EUREX	European Exchange
EURIBOR	Euro InterBank Offered Rate
EWMA	Exponentially Weighted Moving Average
Exc. Kurt.	excess kurtosis
FIEGARCH	Fractionally Integrated Exponential GARCH
FTSE	Financial Times Stock Exchange
GARCH	Generalised Autoregressive Conditional Heteroscedasticity
GED	generalised error distribution
GJR-GARCH	Glosten-Jagannathan-Runkle GARCH
GPH	Geweke-Porter-Hudak
HAR	Heterogeneous Autoregressive
HMAE	mean absolute error adjusted for heteroskedasticity
HMSE	mean square error adjusted for heteroskedasticity
HSIC	heteroskedastic Schwartz information criterion
IGARCH	Integrated GARCH
i.i.d.	independent and identically distributed
ITM	in-the-money
IVAR	integrated variance
IVF	implied volatility function
IVS	implied volatility surface
IVSP	implied volatility spread
IVTSP	implied volatility term structure spread
JB	Jarque-Bera
KKMDB	Karlsruher Kapitalmarktdatenbank

KPSS	Kwiatkowski-Phillips-Schmidt-Shin
LB	Ljung-Box
LHS	left hand side
LIFFE	London International Financial Futures Exchange
LINEX	linear-exponential
LVS	local volatility surface
MA	Moving Average
MAE	mean absolute error
MAPE	mean absolute per cent error
MCS	model confidence set
ME	mean error
MSE	mean square error
NASDAQ	National Association of Securities Dealers Automated Quations
OTC	over-the-counter
OTM	out-of-the-money
PACF	partial autocorrelation function
PDE	partial differential equation
QLIKE	quasi-likelihood
QV	quadratic variation
RMSE	root mean square error
ROB	modified version of the GPH test developed by Robinson (1995b)
RV	realised variance
RVOLA	realised volatility
SD	standard deviation
SIC	Schwartz information criterion
SOFFEX	Swiss Options and Financial Futures Exchange
SPA	superior predictive ability
SV	stochastic volatility
TAR	Threshold Autoregressive
TGARCH	Threshold GARCH

TVSP	total volatility spread
VAR	Vector Autoregressive
VDAX	DAX-Volatilitätsindex
VDAX-New	DAX-Volatilitätsindex New
VIX	CBOE Volatility Index
Xetra	Exchange Electronic Trading

List of Important Symbols

a_t	asset return
A	lag polynomial of order p
b	cost-of-carry
B	lag polynomial of order q
C	call price
CS	cubic spline function
D	dummy variable
dq	Poisson process
dz	Wiener process
E	expectation operator
E^Q	risk-neutral expectation operator
F_t	futures price
F_t^*	forward price
h	bandwidth parameter
h_t	conditional volatility
h_1	bandwidth in the moneyness dimension
h_2	bandwidth in the maturity dimension
I_{t-1}	information set at time $t-1$
j_e	expected percentage change in the asset price due to a jump

J_t	independent identically distributed random variable for the relative change of the asset price in the event of a jump
k	average jump size expressed as a percentage of the asset price
K	strike price
$K_{deep-OTM}^{call}$	strike price of a deep out-of-the-money option
K_h	kernel function
L	lag operator
$m, \widehat{m}$	unknown functional relationship, estimated
M	moneyness
n	number of observations
N	standard normal cumulative distribution function
P	put price
pf	penalizing function
r	risk-free rate
Re	real-valued function
S	asset price
s^2	variance of the jump size
t	time
T	time to maturity
u_t	white noise process
U	dummy variable
V	price of a derivative security
Var	variance operator
$\mathbf{x_t}$	matrix of regressors
ϵ	standardised residuals
ε	error term

ζ	price of volatility risk
η	volatility of volatility
κ	rate of mean reversion
λ	average number of jumps per unit time
μ	drift term
π	number pi
ρ	correlation coefficient
σ	volatility
σ_{IV}	implied volatility
τ	forecast horizon
Ξ	Akaike function
Φ	standard normal cumulative distribution

1. Introduction

1.1. Motivation and Purpose of the Study

Volatility forecasting plays a key role in financial markets. Investors, risk managers, policy makers, regulators, and researchers need volatility forecasts for investment management, security valuation, risk management, and monetary policy. In the following, volatility is defined as the standard deviation of asset returns.

In the field of investment management, volatility is interpreted as a risk measure and used as an input variable for making investment decisions. In a typical investment process, investors define their risk appetite by the level of risk, or volatility, that they are willing to accept. Then, portfolio managers construct portfolios that take these risk preferences into account. Thus, accurate volatility forecasts enable investors to select investment portfolios, that ideally fit their risk-return profiles.

In addition to asset allocation, volatility forecasting is crucial for option pricing. The on-going growth of the global listed derivatives markets, which in 2013 reached 21.64 billion futures and options contracts, emphasises the importance of this area of application.[1] In modern option pricing theory, beginning with Black and Scholes (1973), volatility forecasts are directly plugged into the option pricing formula. Moreover, volatility predictions are needed to hedge portfolios of options. Recently, variance

[1]See Acworth (2014), p. 15.

and volatility swaps have been issued, which allow one to directly invest in volatility as an asset class.[2]

In risk management, the computation of potential portfolio losses, e.g., the value-at-risk, plays a central role. Regulatory authorities, e.g., the European Banking Authority, impose capital requirements that force banks to hold a certain amount of capital to absorb potential future losses. To estimate these potential losses, volatilities and correlations have to be predicted. As many risk models failed in the global financial crisis of 2008, practitioners and academics began to revalidate and revise their risk models, including volatility forecasts. Thus, determining the adequacy of volatility predictions, particularly during episodes of turmoil, is crucial for financial institutions.

Further, governments, central banks, and regulatory authorities also monitor financial market volatility. In particular, they consider implied volatility indices that are regarded as market-based measures of economic uncertainty because studies indicate that volatility can negatively affect the real economy.[3] For example, Bloom (2014) provides evidence that an increase in uncertainty can reduce employment and economic output.[4] Moreover, an empirical paper published by Bekaert et al. (2013) shows that monetary policy affects risk aversion and uncertainty, which both are potentially related to the business cycle.[5] For these reasons, policy makers are interested in the movement of implied volatility indices.

Therefore, a comprehensive overview of volatility prediction models, and a deep understanding of their ability to produce accurate volatility forecasts is an important task in financial market research. This has been subject of a vast number of empirical and theoretical studies over the past few decades.

[2] The pay-out of a volatility swap is equal to the difference between the actual volatility and the predefined contract volatility. See Javaheri et al. (2004), p. 589.

[3] For example, Poon and Granger (2003) refer to the September 11 terrorist attacks and a series of corporate accounting scandals at the beginning of the 21st century.

[4] See Bloom (2014), p. 171.

[5] See Bekaert et al. (2013), p. 771.

The simplest volatility prediction model is to calculate the standard deviation of returns over some historical period. The historical standard deviation is then used as a volatility forecast. While in the past, historical volatility was often used to predict volatility, it has been increasingly replaced by more sophisticated time series models that are able to capture the so-called stylised facts of volatility.

These stylised facts are well documented in the literature. The following features are common to many univariate financial time series:[6]

- *Fat tails.* An important finding is that the distribution of financial returns exhibits fatter tails than the normal distribution.
- *Volatility clustering.* The volatility clustering effect describes the tendency of financial volatility to cluster. This means that a large (small) price change tends to be followed by another large (small) price change.
- *Leverage effect.* An unexpected price drop increases volatility more than an unexpected price increase of equal magnitude.
- *Long memory effect.* Financial market volatility is characterised by long-range dependencies.

Engle (1982) and Bollerslev (1986) developed a new class of volatility models, the Generalised Autoregressive Conditional Heteroscedasticity (GARCH) models that are able to capture the volatility clustering effect. In the basic GARCH model, the conditional variance depends only on own lags and lags of squared innovations. The concept of Engle (1982) and Bollerslev (1986) was extended by Nelson (1991), Glosten et al. (1993), and Zakoian (1994), among others, to model additional empirical characteristics of volatility.[7] Because standard GARCH models fail to capture long memory effects of realised volatilities, Granger and Joyeux (1980) and Hosking (1981) suggested Autoregressive Fractionally Integrated Moving Average (ARFIMA)

[6]See, for example, Brooks (2008), p. 380 and Poon and Granger (2003), p. 482.
[7]In particular, these models capture non-symmetrical dependencies.

processes to parsimoniously model long memory effects.[8] In addition to ARFIMA models, Corsi (2009) suggests a simple AR-type model for realised volatility that is also able to mimic long memory effects.

These time-series models are complemented by implied volatility, which is derived from options prices by applying a particular option pricing model. In this context, implied volatility is interpreted as the market's expectation of the underlying asset's volatility over the remaining lifetime of the option.[9] This interpretation of implied volatility as a market's expectation of future volatility has been criticised because it requires that the assumptions of the applied option pricing model hold.[10] While most early studies used the Black-Scholes (BS) option pricing model to compute implied volatility, alternative models have since been suggested in the literature. In particular, Poteshman (2000), Shu and Zhang (2003), and Chernov (2007) propose stochastic volatility models, whereas Jiang and Tian (2005) recommend model-free implied volatility to forecast US stock market volatility.

However, despite the restrictive and (often) refuted BS assumptions, many studies demonstrate that BS implied volatility provides better volatility forecasts compared with historical volatility models. For example, Poon and Granger (2005) provide a comprehensive literature review and summarise that implied volatility tends to be more appropriate for predicting volatility than historical volatility models, including GARCH models.[11] Thus, some empirical studies show that BS implied volatility provides superior forecasting results, although its model assumptions are violated. In this case, the above-cited theoretical foundation for the use of implied volatility to forecast stock market volatility can no longer be maintained, and the prediction of financial volatility based on BS implied volatility is nothing more than the application of a heuristic rule.

[8] Realised volatility is an ex-post measure for return variation in lieu of the true integrated volatility (see Andersen et al. (2011), p. 221). A detailed definition of realised volatility is given in Section 6.6.1.

[9] See Canina and Figlewski (1993), Mayhew (1995), and Poon and Granger (2003), among many others.

[10] Campbell et al. (1997) note that if the option pricing model does not hold, then the computed implied volatilities are difficult to interpret. See Campbell et al. (1997), p. 378.

[11] See Poon and Granger (2003), pp. 506-507.

Besides the development and application of more suitable option pricing models, papers by Martens and Zein (2004) and Becker et al. (2006) report that long memory models using realised volatility provide good volatility forecasts that can improve implied volatility forecasts by incorporating incremental information. As a consequence, they suggest combined volatility forecasts based on implied volatility and long memory models. Koopman et al. (2005) and Martin et al. (2009) find that long memory models occasionally provide even better prediction results than historical volatility models. In accordance with Martens and Zein (2004), and Becker et al. (2006), they suggest that a combination of individual volatility forecasts from different prediction approaches can improve the performance of volatility forecasts.

In addition to extending the set of forecasting models, studies published by Becker and Clements (2008), Martin et al. (2009), and Martens et al. (2009) use more sophisticated forecast evaluation approaches. Because encompassing regressions only consider individual forecast comparisons, this evaluation approach neglects the comparison of individual forecasts with the complete set of alternative forecasts.[12] By employing the superior predictive ability (SPA) test suggested by Hansen (2005), or the model confidence set (MCS) approach proposed by Hansen et al. (2003), which both allow for a simultaneous comparison of multiple forecasts, the forecast evaluation results of the above studies are not affected by data snooping effects[13].[14]

For the German stock market, the literature presents evidence that DAX implied volatility contains useful information for the prediction of DAX volatility. While Raunig (2006) reports mixed results regarding the relative forecasting performance, recent studies by Muzzioli (2010) and Tallau (2011) suggest that DAX implied volatility provides better volatility forecasts than time series models based on histor-

[12]See Becker et al. (2007), p. 2536.

[13]White (2000) describes data snooping as follows (see White (2000), p. 1097):

> Data snooping occurs when a given set of data is used more than once for purposes of inference or model selection. When such data reuse occurs, there is always the possibility that any satisfactory results obtained may simply be due to chance rather than to any merit inherent in the method yielding the results.

[14]See Hansen et al. (2003), pp. 839-843.

ical returns. Further, Claessen and Mittnik (2002) indicate that combined volatility forecasts using the information from implied volatility and historical returns are a reasonable complement to individual forecasts. In addition, Lazarov (2004) presents the interesting result that the forecasting performance of ARFIMA models is similar to that of DAX implied volatility. In summary, all these studies of the German stock market examine a subset of forecasting models, but do not provide a comprehensive comparison of volatility prediction models. Further, because the above studies concerning the performance of DAX volatility prediction models employ encompassing regressions, this is to my knowledge the first study to evaluate DAX volatility forecasts based on the MCS method and DAX realised volatility.[15]

While a variety of extensive studies on the forecasting performance of implied volatility computed from various option pricing models, time series models, and combinations thereof have been published for the US stock market, a similar study for the German stock market that considers all these forecasting models does not exist. In addition, a forecast evaluation approach that controls for data snooping effects has not been applied to compare the prediction results of these models for the German stock market. The intent of this study is to close these research gaps and to provide information to investment and risk managers regarding which forecasting method delivers superior DAX volatility forecasts.

The empirical analysis is based on data that contain all recorded transactions of DAX options and DAX futures traded on the EUREX from January 2002 to December 2009. To select an appropriate time series model for the prediction of DAX volatility, the time series features of DAX returns and realised volatilities are investigated in this thesis. Further, this study presents a detailed analysis of the characteristics of the DAX implied volatility surface (IVS) because the results will provide information for selecting an adequate option pricing model. In particular, the DAX IVS is investigated for three different subsamples because different volatil-

[15]Further, the MCS approach allows to select the best forecasting models from a range of models, and it is not necessary to define any specific benchmark model to determine the MCS. Additionally, predictions can be compared based on different loss functions.

ity regimes occurred during the sample period. If the DAX IVS exhibits certain regularities during silent and/or turbulent market periods, the selected option pricing model should be flexible enough to capture these effects. To my knowledge, this is the first comprehensive analysis of the impact of the financial crisis of 2008 on the DAX IVS.

Due to the discrepancies between the observed patterns of the DAX IVS and the assumptions of the BS model, which provide evidence against the suitability of the BS model for pricing DAX options, the methodologies of alternative option pricing models are presented. In addition, given the time series features of the DAX returns and realised volatilities documented by this study, the methodology of appropriate time series models is described.

After the introduction of the theory underlying the forecasting approaches, a literature review presents the empirical results of selected studies that compare these approaches' forecasting performance for the US stock market. As such a broad and deep discussion does not exist for the German stock market, the findings of these papers provide useful information for the empirical analysis performed for German stock market volatility.

The volatility prediction models employed in this study to forecast DAX volatility are selected based on the results of these empirical studies, the general features of the forecasting models, and the analysis of the considered DAX time series. Within the class of time series models, the GARCH, the Exponential GARCH (EGARCH), the ARFIMA, and the Heterogeneous Autoregressive (HAR) model are chosen to fit the DAX return and realised volatility series. Additionally, the Britten-Jones and Neuberger (2000) approach is applied to produce DAX implied volatility forecasts because it is based on a broader information set than the BS model. Moreover, recent studies report promising empirical results with respect to the forecasting performance of this approach. Finally, the BS model is employed as a benchmark model in this study.

As the empirical analysis in this study demonstrates that DAX volatility changes considerably over the long sample period, it investigates whether structural breaks induce long memory effects. The effects are separately analysed by performing different structural break tests for the prediction models. A discussion of the impact on the applied forecasting methodology, and how it is accounted for, is also presented.

Based on the MCS approach, the DAX volatility forecasts are separately evaluated for the full sample and the subperiod that excludes the two most volatile months of the financial crisis. Because the objective of this work is to provide information to investment and risk managers regarding which forecasting method delivers superior DAX volatility forecasts, the volatilities are predicted for one day, two weeks, and one month. Finally, the evaluation results are compared with previous findings in the literature for each forecast horizon.

Overall, this study provides a comprehensive comparison of different forecasting approaches for the German stock market, which yet does not exist. Additionally, this thesis presents the first application of the MCS approach to evaluate DAX volatility forecasts based on high-frequency data. Furthermore, the effects of the 2008 financial crisis on the prediction of DAX volatility, that are not considered in the literature, are analysed.

1.2. Overview of the Thesis

After defining the purpose of the study, Chapter 2 outlines the basic concept of using implied volatilities to predict volatility. Because the BS option pricing model is considered a cornerstone in the history of pricing contingent claims and is often applied as a reference model in empirical studies, it is introduced in this Chapter. Subsequently, a critical review of the use of BS implied volatilities as volatility forecasts is presented. Further, some stylised facts of BS implied volatilities documented in the literature are described. Finally, selected potential explanatory approaches for the observed BS implied volatility patterns are discussed.

Having presented the pricing biases of the BS model that are documented in the literature, the aim of Chapter 3 is to investigate the (mis)pricing behaviour of the BS model for the German stock market. In particular, this study considers DAX options traded on the EUREX from January 2002 to December 2009. To analyse BS implied volatilities across moneyness and maturity, it is necessary to construct a smooth IVS. Thus, the basic concepts of two general smoothing approaches are discussed, and the choice of the approach employed in this study is explained. Thereafter, the characteristics of the DAX BS IVS are described and compared to the existing literature. Moreover, Chapter 3 presents the underlying data set and its preparation.

Because the empirical analysis of the DAX BS IVS demonstrates that some of the BS model assumptions are violated, a range of alternative option pricing models is presented in Chapter 4. To capture the stylised features of the IVS, these models relax some of the BS assumptions. In addition to stochastic volatility and mixed jump-diffusion models, the concept of model-free implied volatility developed by Britten-Jones and Neuberger (2000) is presented. The introduction of this concept is completed by a critical review that considers some additional assumptions necessary for its implementation. The ability of each model class to reproduce the observed DAX IVS is discussed at the end of each Section. In addition to the option pricing models, selected time series models are described that are used in this study to predict DAX volatility.

Chapter 5 presents a literature review of empirical studies comparing the volatility forecasting performance of implied volatility and time series models. Because most early studies use encompassing regressions to evaluate volatility forecasts, the first Section of Chapter 5 introduces this evaluation method. The second Section reviews selected papers on predicting US stock market volatility, as these articles contain broad and intensive discussions of the US stock market. The following Section introduces empirical studies on the predictive ability of implied volatility and time series models for German stock market volatility. The final Section explains the choice of the volatility prediction models used in this study to forecast DAX volatility.

Chapter 6 focuses on the generation and evaluation of the DAX volatility forecasts. Due to the characteristics of the DAX return and volatility series, the GARCH, the EGARCH, the ARFIMA, and the HAR models are estimated. Then, information criteria are used to select the most appropriate models. Subsequently, the effects of structural breaks on long memory effects are analysed. Afterwards Section 6.6 explains the choice of the employed volatility proxy, realised volatility, and describes its calculation for the DAX return series. Further, this Chapter contains a brief overview of forecast evaluation techniques and arguments for the selection of the MCS approach applied in this thesis. In the following, the DAX volatility predictions based on the above models are presented for different forecast horizons. In addition to the individual forecasts, this study also considers combined forecasts because some forecast combinations have been found to outperform individual forecasts. Finally, the prediction results are evaluated by using the MCS approach and compared with the previous findings in the literature.

The final Chapter summarises the results, provides recommendations for predicting German stock market volatility, and presents an outlook on future research, including possible extensions of this thesis.

2. The Concept of Implied Volatility

Two approaches for predicting financial volatility have been suggested in the literature. The first of these involves the generation of volatility forecasts based on time series models. In the second approach, implied volatilities, which are derived from option prices, can be used to forecast financial volatility.[1] First, this Chapter outlines the basic concept of using implied volatilities to predict volatility.[2] Next, the BS option pricing model is introduced[3], which is applied in this study to calculate implied volatilities from DAX option prices. Finally, the stylised facts of implied volatilities that have been documented in the literature are presented and several corresponding explanatory approaches are discussed.

2.1. The Basic Concept of Using Implied Volatilities to Forecast Volatility

An option is a derivative security, the price of which depends on the future development of the underlying asset price. Therefore, option pricing models generally specify a stochastic process to model the price of the underlying asset. Asset volatility is typically one of the main parameters in this process. To compute an option price based on an option pricing model, the volatility parameter has to estimated and plugged into an option pricing formula. Conversely, using an option pricing

[1]See Poon and Granger (2003), p. 482.
[2]See Chapter 4 for an introduction to time series models.
[3]See Black and Scholes (1973).

model, e.g., the BS model, volatility can also be deduced from the option price. This volatility is called *implied volatility*.[4]

Implied volatility is widely interpreted as a market's expectation of the underlying asset's volatility over the remaining lifetime of the option, as it is derived from the market price.[5] Thus, it is regarded as a "forward-looking" volatility estimate of the return on the underlying asset,[6] which should provide the market's best volatility forecast over the option's maturity.[7] Moreover, under the assumption of market efficiency, implied volatility should provide an informationally efficient forecast of volatility that also contains information on the historical returns of the underlying asset.[8,9]

Latané and Rendleman (1976) provide the first study on the forecasting ability of implied volatilities. They investigate individual stock options traded on the Chicago Board Options Exchange (CBOE) in 1974 and report that a weighted average of implied volatilities is a better predictor of volatility than the standard deviation based on historical returns.[10] Although numerous articles have been published on the forecasting performance of implied volatilities and time series models, the debate over which approach delivers better volatility forecasts persists. A comprehensive literature overview of this discussion is provided in Chapter 6.

However, Campbell et al. (1997) criticise the interpretation of implied volatility as a market's expectation of future volatility. They argue that implied volatility, which is calculated based on a specific option pricing model, is inseparably related to the model-implicit dynamics of the underlying asset price. Thus, interpreting implied volatility as a market's prediction of future volatility requires that the option pricing model holds. If the option pricing model does not hold, then the computed implied

[4] See Rouah and Vainberg (2007), p. 322.
[5] See Canina and Figlewski (1993), Mayhew (1995), and Poon and Granger (2003) among others.
[6] See Rouah and Vainberg (2007), p. 304.
[7] See Ederington and Guan (2005), p. 1429.
[8] See for example Christensen and Prabhala (1998).
[9] The literature on the informational content of option prices is discussed in Chapter 5.
[10] See Latané and Rendleman (1976), p. 381.

volatilities are difficult to interpret.[11] Thus, predictive ability tests evaluating implied volatility forecasts are joint tests of predictive ability and the applied option pricing model.[12]

However, as many studies have demonstrated that implied volatilities exhibit superior predictive ability for various options markets, this study investigates their forecasting performance. To account for the argument advanced by Campbell et al. (1997), two different approaches are used to derive implied volatility from option prices. In particular, the approach developed by Britten-Jones and Neuberger (2000) is used to calculate mode-free implied volatility, which does not require the specification of a particular process for the price of the underlying asset.[13] Nonetheless, the above critique must be kept in mind when the results of the predictive ability tests for implied volatilities are discussed. In the following the most popular option pricing model is presented, namely the BS model, which is used in this study to calculate implied volatilities from option prices.

2.2. The Black-Scholes Model

The development of the BS option pricing model by Black and Scholes (1973) and further by Merton (1973) marks a breakthrough in financial theory. They show that under certain conditions, markets are complete and contingent claim valuation is preference-free. As different studies demonstrate that its assumptions are rather restrictive, the model has been extended in the subsequent literature, and there are currently a large number of refined models available.[14] Despite the extensions, the BS model is considered a cornerstone of pricing contingent claims and is used as a reference model in empirical studies.[15] For this reason, the Black-Scholes par-

[11] See Campbell et al. (1997), p. 378.
[12] See Jiang and Tian (2005), p. 1306.
[13] Alternatively, they derive a condition that characterises all continuous price processes that are consistent with current option prices. See Britten-Jones and Neuberger (2000), p. 839.
[14] See Fengler (2004), p. 9.
[15] See Rebonato (2004), p. 168.

tial differential equation (BS PDE), which under certain assumptions describes the option's equilibrium price path is derived in the following. The BS option pricing formula is presented on the BS PDE.

To develop the BS option pricing model, Black and Scholes (1973) rely on several assumptions. First, they assume that market participants can trade continuously in a frictionless market where no arbitrage possibilities exist.[16] Further, the underlying asset pays no dividends, assets are divisible, and short selling is allowed. In addition, investors can lend or borrow without restrictions at the same riskless rate of interest. Moreover, the risk-free interest rate is known and constant over time. While some of the assumptions are not necessary to derive the option pricing model, the following assumption regarding the dynamics of the asset price is essential.[17] Black and Scholes (1973) assume that the asset price follows a geometric Brownian motion

$$dS = \mu S dt + \sigma S dz \tag{2.1}$$

where S denotes the underlying asset price, μ the instantaneous drift, σ the instantaneous volatility, and dz a Wiener process.[18]

Suppose that V represents the price of an option or other derivative security, the price of which exclusively depends on S and time t. From Itô's lemma, it follows that V can be written as

$$dV = \left(\frac{\partial V}{\partial S} \mu S + \frac{\partial V}{\partial t} + \frac{1}{2} \frac{\partial^2 V}{\partial S^2} \sigma^2 S^2 \right) dt + \frac{\partial V}{\partial S} \sigma S dz. \tag{2.2}$$

For a brief time interval Δt the discrete versions of (2.1) and (2.2) are given by[19]

$$\Delta S = \mu S \Delta t + \sigma S \Delta z \tag{2.3}$$

[16] In a frictionless market no transaction costs and or taxes exist.
[17] Black and Scholes (1973) formulate certain assumptions for expositional convenience.
[18] See ibid., p. 640.
[19] See Hull (2006), p. 291.

and

$$\Delta V = \left(\frac{\partial V}{\partial S} \mu S + \frac{\partial V}{\partial t} + \frac{1}{2} \frac{\partial^2 V}{\partial S^2} \sigma^2 S^2 \right) \Delta t + \frac{\partial V}{\partial S} \sigma S \Delta z. \tag{2.4}$$

Based on the underlying asset and the derivative, a portfolio with the value

$$\Pi = -V + \frac{\partial V}{\partial S} S \tag{2.5}$$

is constructed.[20] In a brief time interval Δt, the portfolio value changes by

$$\Delta\Pi = -\Delta V + \frac{\partial V}{\partial S} \Delta S. \tag{2.6}$$

Replacing ΔV and ΔS in (2.6) with (2.3) and (2.4) yields:[21]

$$\Delta\Pi = -\frac{\partial V}{\partial t} \Delta t - \frac{1}{2} \frac{\partial^2 V}{\partial S^2} \sigma^2 S^2 \Delta t. \tag{2.7}$$

Because the Wiener processes in (2.3) and (2.4) are identical, they are eliminated in (2.7). It follows that the portfolio in the time interval Δt is riskless. Thus, the portfolio return must be equal to the risk-free rate, which can be expressed by

$$\Delta\Pi = \Pi r \Delta t \tag{2.8}$$

where r is the risk-free rate.[22] It should be noted that the portfolio is only riskless over an infinitesimal time interval. To ensure that the portfolio is riskless over time, a dynamic hedging strategy is necessary (e.g., delta hedging).[23]

Substituting equations (2.5) and (2.7) into (2.8) yields the BS PDE

$$\frac{\partial V}{\partial t} + rS \frac{\partial V}{\partial S} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} = rV. \tag{2.9}$$

[20] The portfolio consists of a short position in the derivative and a long position in the amount of $\partial V / \partial S$ in the underlying asset.

[21] See Wilmott et al. (1993), p. 43.

[22] Otherwise riskless arbitrage opportunities would exist which is ruled out by the above assumptions. See Wilmott et al. (1993), pp. 43-44.

[23] See Hull (2006), p. 292.

Under the above assumptions the price of any derivative security must satisfy the BS PDE.[24] The solution of the BS PDE depends on the considered derivative which is specified by its boundary conditions. For instance, for a European call option the boundary condition is

$$V = \max(S - K, 0) \quad \text{when } t = T \tag{2.10}$$

where K represents the strike price and T the time to maturity. Based on this final condition, a unique solution for the BS PDE can be derived.[25] In the following the solution of the BS PDE for a European call option is presented. For a detailed derivation see, for instance, Ekstrand (2011).

The solution of the BS PDE for a European call option, which is also called the *BS formula* is given by

$$C(S,t) = SN(d_1) - Ke^{-r(T-t)}N(d_2) \tag{2.11}$$

with

$$d_1 = \frac{\ln(S/K) + (r + \sigma^2/2)(T-t)}{\sigma\sqrt{T-t}} \tag{2.12}$$

$$d_2 = \frac{\ln(S/K) + (r - \sigma^2/2)(T-t)}{\sigma\sqrt{T-t}} = d_1 - \sigma\sqrt{T-t} \tag{2.13}$$

where $C(\cdot)$ denotes the price of a European call option and $N(\cdot)$ is the standard normal cumulative distribution function.[26] The price of a European put option $P(\cdot)$ can be calculated based on the put-call parity by[27]

$$P(S,t) = C(S,t) + Ke^{-r(T-t)} - S. \tag{2.14}$$

[24]See Wilmott et al. (1993), p. 44.
[25]See Joshi (2003), p. 105.
[26]See Wilmott et al. (1993), p. 49.
[27]See Hull (2006), p. 212.

The assumptions of the BS model have been intensively criticised in the literature.[28] In practice, a frictionless market does not exist, a continuous hedge without transaction costs is impossible, and the asset price does not follow a geometric Brownian motion. Deviations from these assumptions affect the option price and therefore the implied volatility. Thus, whether and in particular to what extent these violations of the BS assumptions occur can be investigated based on implied volatilities. For this reason, the next Section presents the methodology for deriving BS implied volatilities from option prices.

2.3. Calculation Methodology of Black-Scholes Implied Volatilities

According to the BS formula, the option price depends on the current time, the level and volatility of the underlying asset price, the interest rate, the strike price, and the maturity date. Except for volatility, all parameters are determined by the contract specification or can be directly observed in the market. As these parameters are fixed, the BS formula defines a one-to-one relationship between the option price and volatility. This, the volatility implied by the market price can be determined by the inverse of the BS formula.[29]

Formally, given an observed market price $V_{obs}(K,T)$ of an European option with strike price K and time to maturity T, the BS implied volatility σ_{IV} is defined as the value of volatility in the BS formula for which the BS option price V_{BS} is equal to the market price:[30]

$$V_{obs}(K,T) = V_{BS}(\sigma_{IV}, K, T). \tag{2.15}$$

[28] See, for instance, Gourieroux and Jasiak (2001a), pp. 321-323, Musiela and Rutkowski (2005), p. 113, and Chriss (1997), pp. 200-204.
[29] See Ekstrand (2011), p. 30.
[30] See Rouah and Vainberg (2007), p. 305.

The existence of a unique solution is ensured, as the BS formula is monotonically increasing in volatility.[31]

The implied volatility cannot be extracted from the BS formula analytically.[32] Instead, it can be computed numerically by finding the root of the objective function

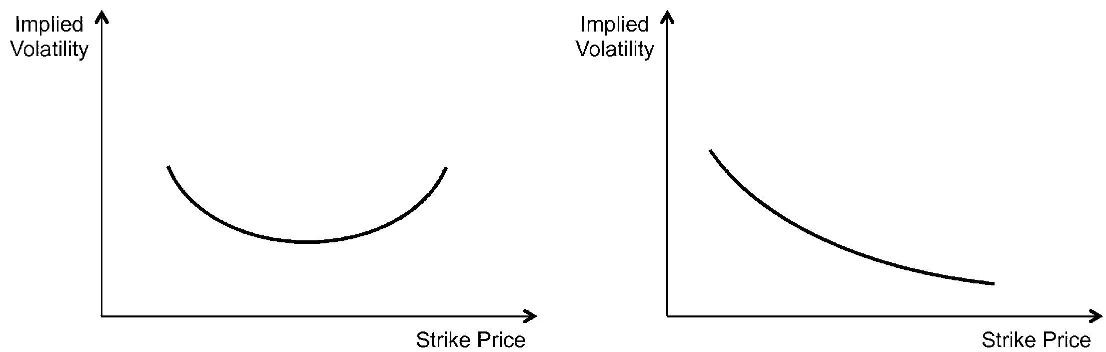
$$f(\sigma) = V_{BS}(\sigma, K, T) - V_{obs}(K, T) \tag{2.16}$$

such that $f(\sigma_{IV}) = 0$.[33] The optimisation problem can be solved for instance by the Newton-Raphson method or the bisection method. It is well known that the Newton-Raphson algorithm is quite sensitive to the initial volatility value, which can lead to unfavourable solutions.[34] Further, it requires that the derivative of the option price with respect to the volatility parameter (vega) is known or can be approximated numerically. In contrast, the bisection method avoids the need for knowledge of vega, as it is based on a simple interpolation method.[35] Due to the Intermediate Value Theorem, the algorithm always finds one root of the above objective function for volatility intervals in which the objective function changes its sign.[36] A disadvantage of the bisection method is that it is not as fast as the Newton-Raphson method. As the Newton-Raphson algorithm can diverge from the root and the speed of the bisection algorithm for the computation of DAX implied volatilities is acceptable, this study employs the bisection algorithm to find the roots of the above objective function.[37]

While, in theory, the BS implied volatilities of all options on the same underlying asset should be identical, in practice, they are not. It is well documented in

[31] See Cont and da Fonseca (2002), p. 47.
[32] For the special case of at-the-money (ATM) options, Brenner and Subrahmanyan (1988) demonstrated that implied volatility can be calculated by a simple approximation formula derived from the BS model.
[33] See Rouah and Vainberg (2007), p. 305.
[34] See ibid., p. 307.
[35] See Haug (2007), p. 455.
[36] See Rouah and Vainberg (2007), p. 9.
[37] To assess the effect of the selected algorithm on the resulting implied volatilities, the DAX implied volatilities are calculated for a subsample based on both algorithms. The results indicated that the differences between the implied volatilities obtained by the Newton-Raphson and the bisection algorithm are very small.

the literature that BS implied volatilities are not constant across strike prices and maturities. Rather, for many options markets, systematic patterns of BS implied volatilities across strike prices and across maturities have been observed. The plot of implied volatilities of options with the same maturity but different strike prices (or moneyness levels) is typically U-shaped. This well-known phenomenon is referred to as the *volatility smile*. If the shape of the volatility smile is asymmetric, it is called *volatility skew*. Figure 2.1 depicts two types of volatility smiles. The functional relationship between implied volatility and strike price/moneyness is also called the implied volatility function (IVF).[38]



(a) Smile (b) Skew

Figure 2.1.: Types of volatility smiles
Source: own illustration.

Alternatively, if the implied volatilities of options of the same strike price (or moneyness level) but different maturities are considered, the implied volatility moves with increasing maturity towards long-term implied volatility. This is called the *volatility term structure*.[39] The combined analysis of the relationship between implied volatilities and strike prices (volatility smile) and maturities (volatility term structure) is based on the so-called *implied volatility surface* (or *volatility surface*).

[38]The term *volatility smile* is synonymously used for the implied volatility function and its plot.
[39]See Alexander (2008), p. 227.

An IVS depicts the implied volatilities for different strike prices and maturities.[40] The BS IVS[41] at time t is defined as

$$\sigma^{iv} : (t, K, T) \rightarrow \sigma^{iv}(K, T). \tag{2.17}$$

Thus, when the IVS can be fully identified at time t, this means that all (vanilla) call and put option prices are known at t.[42] Before presenting findings from the empirical literature on the volatility smile and term structure, the next Section discusses the use of implied volatilities as volatility forecasts.

2.4. A Critical Review of Using Implied Volatilities as Volatility Forecasts

The interpretation of BS implied volatility as the market's expectation of the (constant) volatility of the underlying asset over the lifetime of the option is based on the validity of the BS model. In particular, the BS model has been criticised for its unrealistic assumption of constant volatility.[43] Feinstein (1989) shows that this assumption can be relaxed. He demonstrates that the BS implied volatility from a near-expiration ATM call option yields an unbiased forecast of the average volatility over the remaining life of the option when volatility is stochastic and uncorrelated with aggregate consumption.[44] The assumption of zero correlation between volatility changes and aggregate consumption changes ensures that volatility risk is unpriced in the market. However, the empirical results of Lamoureux and Lastrapes (1993) call this assumption into question, as they report that the market price of volatility risk for individual US stocks is nonzero and time varying. Thus, Feinstein' s (1989) finding cannot be used as a general argument in favour of the BS model. For this

[40] See Hull (2006), p. 382.
[41] The term *volatility surface* is used for the preceding function and its graphical representation.
[42] See Cont and da Fonseca (2002), p. 45.
[43] See, for instance, Gourieroux and Jasiak (2001a), p. 279.
[44] The Hull and White (1987) option pricing model provides the theoretical framework for this argument.

reason, advanced option pricing models, e.g., the stochastic volatility model suggested by Heston (1993), have been developed to take time-varying volatility into account.

Despite the unrealistic assumptions of the BS model and the availability of stochastic volatility models, option traders typically quote prices in terms of BS implied volatilities. By expressing option prices as BS implied volatilities, traders seek to control for different strike prices and maturities.[45] Campbell et al. (1997) argue that this only reflects the popularity of the BS formula as a heuristic, but has no economic implications. Further, they note that option traders quoting prices using BS implied volatilities does not necessarily imply that they calculate their prices based on the BS formula. They conclude that due to the one-to-one relationship between BS implied volatilities and option prices, both pricing measures cover the same information.[46] Thus, if the BS model does not hold, the use of BS implied volatilities as volatility forecasts corresponds to the application of a heuristic rule. However, even if one agrees with this interpretation of BS implied volatility, it is an interesting research topic to compare its forecasting ability with those of alternative prediction approaches. The next Section reports some well-known stylised facts of implied volatilities, including the volatility smile and the volatility term structure, which have been documented in the literature.

2.5. Stylised Facts of Implied Volatilities

This Section begins with a review of the discussion on the volatility smile effect. The discussion covers two basic forms of the volatility smile and presents some important empirical results for different options markets. Subsequently, the term structure of implied volatility is described. Finally, the time series properties of implied volatilities that have been observed for some selected options markets are presented.

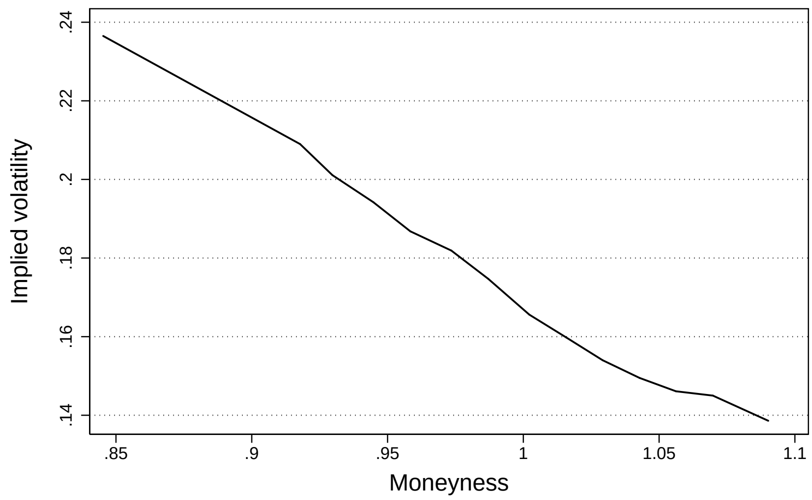
[45]See Hafner (2004), p. 37.
[46]See Campbell et al. (1997), p. 379.

2.5.1. The Smile Effect

From Smiles and Skews

The first two articles documenting the systematic pattern of BS pricing errors across strike prices and maturities are Black (1975) and MacBeth and Merville (1979). Using CBOE prices for the period from 1973 to 1975, Black (1975) reports that the actual market prices of out-of-the-money (in-the-money) options tend to be higher (lower) than the values given by the BS formula. He suggests different explanations for this pattern including time-varying volatility, tax factors, speculative profits, and leverage restrictions.[47] In contrast, MacBeth and Merville (1979) find that implied volatilities of CBOE options tend to increase with decreasing strike prices for the period from 1975 to 1976. Rubinstein (1985) analyses trades and quotes on the 30 most active option classes for individual stocks in the CBOE from August 1976 to August 1978 and finds that the implied volatilities of short-term out-of-the-money (OTM) calls are higher than for other calls.[48] By considering different time periods, he demonstrates that the sign of the price differences between market prices and BS values changes over time. However, he notes that while the option price differences are indeed statistically significant, their economic significance is questionable.[49] Moreover, Mixon (2009), who investigates empirical regularities of implied volatilities in the 19th century and the 21st century, provides early evidence in favour of the existence of an implied volatility skew in the 19th century.[50] Thus, the findings of Black (1975), Rubinstein (1985), Mixon (2009), and others show that the volatility smile phenomenon already existed at least in a weak form in some options markets before the stock market crash in 1987, which marks a decisive turning point for volatility smiles.[51]

[47] See Black (1975), pp. 64-65.
[48] See Rubinstein (1985), p. 474.
[49] See ibid., p. 478.
[50] See Mixon (2009), p. 172.
[51] The dramatic decline in the stock market on October 19th and 20th, 1987 of more than 20% was the greatest decline since 1929. The crash of 1987 was preceded by an extraordinary increase of 42% in 1987.



Source: See Rubinstein (1994), p. 776.
Note: Implied volatilities based on S&P 500 index options from January 2nd, 1990 at 10:00 a.m.

Figure 2.2.: Typical S&P 500 post-crash volatility smile

Prior the crash of 1987, implied volatilities in stock options markets typically formed a symmetric smile pattern when plotted against strike price or moneyness. After the crash of 1987, the shape of implied volatilities for most stock index options markets more resembles a skew, where implied volatilities decrease monotonically as the strike price rises.[52,53] Figure 2.2 depicts a typical downward sloping post-crash smile for S&P 500 index options. The change in the form of the volatility smile towards a skew shape implies that after the crash of 1987, market participants have paid higher prices for OTM put and in-the-money (ITM) call options relative to other options. Rubinstein (1994) argues that the overpricing of put options is induced by an excess demand for put options, which provide portfolio insurance against market downturns. The excess demand for put options reflects investors' concerns regarding another stock market crash. Rubinstein (1994) terms this phenomenon "crash-o-phobia".[54] The volatility skew phenomenon did not disappear following the crash. This also implies that the implicit distribution of option prices has shifted

[52]See, for example, Poon and Granger (2003), p. 487 and Cont and Tankov (2004), p. 10.
[53]In addition to this permanent effect, implied volatility remained at a high level for several months after the crash (implied volatility more than doubled), but had returned to a pre-crash level by March 1988.
[54]See Rubinstein (1994), p. 775.

from a widely symmetric and positively skewed distribution to a substantially negatively skewed distribution.[55] As Rubinstein (1994) highlights, the crash of 1987 permanently changed market participants' perceptions and pricing mechanisms for stock index options. For this reason, in the next Chapter, this study also investigates whether the 2008 financial crisis affects the volatility smile of DAX options. The following Section provides an overview of the empirical results concerning volatility smiles across different options markets.

Empirical Evidence for Volatility Smiles Across Different Options Markets

The volatility smile effect has been observed for many options markets.[56] Analysing daily over-the-counter (OTC) implied volatility quotes on 12 major equity indices (i.e., CAC (France), DAX (Germany), FTSE (United Kingdom), HSI (Hong Kong), NKY (Japan), and SPX (United States)) from June 1995 to May 2005, Foresi and Wu (2005) confirm the existence of a heavily skewed implied volatility pattern for all indices examined. Interestingly, they find that the markets differ more in the level of implied volatility than in the skewness of the volatility smile. Overall, they conclude that the volatility skew is not a local observation, but rather a worldwide phenomenon.[57]

In a broad study, Tompkins (2001) considers 16 options markets with respect to stock indices, bonds, exchange rates, and forward deposits over long time periods and compares the regularities of the IVS across the different markets. His sample comprises, for most markets, option closing prices over ten years beginning in the mid-1980s and ending in the mid-1990s.[58] Overall, he finds that the shapes of the implied volatilities, which are smoothed based on a quadratic regression, exhibit

[55] See Bates (2000), p. 182.
[56] For a comprehensive overview of stock, bond and exchange rate markets, see, e.g., Rebonato (2004).
[57] See Foresi and Wu (2005), pp. 11-13.
[58] To avoid the problem of non-synchronous trading, Tompkins (2001) only examines OTM put and call options with maturities up to 90 days. To control for level effects, which according to Dumas et al. (1998) contain no exploitable information on future levels of implied volatility, he uses standardised implied volatilities.

similar patterns for options within the same asset class. First, the smoothed implied volatilities of short-term stock options on the S&P 500, the FTSE, the Nikkei, and the DAX index are "U-shaped" across moneyness. Second, the volatility smiles for options with 90 days to maturity generally exhibit a comparatively linear skewed form.[59] Moreover, the regression results support the findings of Rubinstein (1994) and others that the negative volatility skew of S&P 500 options is related to the 1987 stock market crash. However, he reports that a second shock in 1989 also contributes to the negative skewness of the S&P 500 volatility smile.[60] Furthermore, he mentions that the IVS in all considered stock index options markets becomes flatter when the level of implied volatility of ATM options increases in the markets.[61]

In addition to the above basic facts regarding stock volatility smiles, Rebonato (2004) adds that the smile is generally much more pronounced at short maturities and flattens out at longer maturities. Furthermore, the smile of OTM puts is typically steeper than the smile of OTM calls. In some cases, the smile completely disappears for OTM calls. Moreover, during periods of high volatility, Rebonato (2004) notes that the asymmetry of the smile usually tends to increase.[62] Next, the empirical results concerning the volatility smile for the German options market are presented.

Empirical Studies of the DAX Volatility Smile

Few empirical findings regarding the existence of a volatility smile in the German options market before the 1987 crash are provided in the literature. For instance, Trautmann (1990) describes systematic pricing biases in the BS model. He considers pricing differences between market prices and BS values of individual stock options

[59] Tompkins (2001) reports comparable findings for interest rate options (US Treasury Bonds, UK Gilts, German Bundesanleihen, and Italian Government Bonds). The volatility smiles of foreign exchange rate options (US Dollar/Deutsche Mark, US Dollar/British Pound, US Dollar/Japanese Yen and US Dollar/Swiss Franc) exhibit the most similarities. See Tompkins (2001), p. 204.

[60] While both shocks also change the smile of FTSE 100 index options, the DAX volatility smile is unaffected.

[61] See Tompkins (2001), pp. 200-218.

[62] See Rebonato (2004), p. 206.

traded on the Frankfurter Optionsbörse from 1983 to 1987. Based on these data, he demonstrates that the BS model underprices deep OTM call options on individual stocks with maturities beyond four weeks.[63] Thus, the results indicate that the BS assumption of constant volatility had been violated for individual stock options traded on the Frankfurter Optionsbörse before the stock market crash of 1987.[64]

Evidence for the existence of a post-crash DAX volatility smile is reported by Beinert and Trautmann (1995), Ripper and Günzel (1997), Herrmann (1999), and Bolek (1999), among others. Beinert and Trautmann (1995) use transaction prices for the most liquid call options on individual stocks traded on the Deutsche Terminbörse (DTB) from 1990 to 1991. In particular, they demonstrate the volatility smile of short-term options has a U-shaped form. In summary, they conclude that the smile pattern is typical of options traded on the DTB, but did not hold in every case.[65] Ripper and Günzel (1997) estimate the IVS based on a regression using daily settlement prices on DAX options for the years 1995 to 1996 listed on the DTB. The results indicate a volatility smile for short-term options and a skew for long-term options (with maturities beginning at three months).[66] Herrmann (1999) confirms the results of Ripper and Günzel (1997) with respect to the existence of a volatility smile for short-term options. His analysis is based on transaction data of DAX options traded on the DTB from 1992 to 1997. Additionally, he finds that the implied volatilities of ITM and OTM options in the same moneyness class typically decline with increasing maturity. In contrast, he reports increasing implied volatilities for ATM options if the time to maturity is extended. Further, the highest implied volatilities are recorded for deep ITM calls and deep ITM puts. Moreover, he reports that DAX puts have higher implied volatilities than DAX calls of the same moneyness/maturity class.[67] Bolek's (1999) study is based on the closing prices of DAX options traded on the DTB in the 2nd half of 1995. His findings also

[63] See Trautmann (1990), p. 95.
[64] As was the case for the US stock market, the DAX declined dramatically in 1987 (approximately 22% from September 1987 to October 1987).
[65] See Beinert and Trautmann (1995), p. 13.
[66] See Ripper and Günzel (1997), p. 475.
[67] See Herrmann (1999), p. 185.

support the existence of a DAX volatility smile, especially for short-term options.[68] The minimum of the smile of short-term DAX options is located ATM. For long-term DAX calls, the minimum of the smile is OTM. In this case, the smile nearly disappears. In contrast, the minimum of the DAX volatility smile for long-term put options is ITM and the smile is less pronounced. By plotting the volatility smile implied by DAX calls and puts for successive trading days, he finds evidence that the smile changes over time.[69]

Subsequently, Tompkins (2001) and Wallmeier (2003) also examine the systematic pattern of DAX implied volatilities.[70] Tompkins (2001) analyses DAX option closing prices for the time period from January 1992 to December 1996. Using the quadratic regression approach suggested by Shimko (1993), he finds that the volatility smile is negatively skewed. Further, the IVF was more skewed for short-term options than for options with longer maturities.[71] Wallmeier (2003) investigates transaction prices of DAX options for the period from 1995 to 2000. His results also reveal the existence of a DAX volatility skew rather than a smile. Moreover, he reports that the magnitude of the skew tends to decrease as the level of ATM implied volatility rises. He concludes that in the sample period, DAX implied volatilities differ considerably across strike prices and maturities. Therefore, he notes that the BS assumption of constant volatility is seriously violated.[72]

More recent evidence concerning DAX implied volatility smiles is provided by Brunner and Hafner (2003), Fengler et al. (2003), Hafner (2004), Fengler et al. (2007), Detlefsen (2007), Schnellen (2007), and Brüggemann et al. (2008).[73] Walter (2008) considers the volatility smile of individual German stocks. The next Section describes empirical regularities of the volatility term structure

[68]Bolek (1999) defines short-term options as options with maturities of up to 15 trading days.
[69]See ibid., pp. 122-128.
[70]In particular, both studies consider the factors influencing the IVS. The results are presented in the last Section of this Chapter.
[71]See Tompkins (2001), p. 204.
[72]See Wallmeier (2003), pp. 190-208.
[73]The studies typically investigate the time series behaviour of the IVS.

2.5.2. The Volatility Term Structure

Basic Volatility Term Structures

In addition to providing extensive evidence of the volatility smile, the term structure of implied volatility is intensively discussed in literature. In Chapter 2.3, the term structure of volatility was defined as the relationship between implied volatility and time to maturity for a given strike price (or moneyness level).[74] The constant volatility assumption of the BS model implies a flat term structure. Therefore, the implied volatilities of short-term and long-term options should be identical. In contrast, various authors report empirical evidence of a non-flat volatility term structure for many options markets. Similar to the yield curve, the volatility term structure is interpreted as the market expectation of future volatility changes. Figure 2.3 depicts the two basic profiles of the volatility term structure. A downward (upward) sloping volatility term structure indicates that market participants expect decreasing (increasing) future short-term implied volatilities.[75,76] As in interest rate markets, an upward sloping volatility term structure is called *normal*. Alternatively, if the term structure takes a negative slope, it is referred to as *inverse*.[77] Fengler (2012) adds that a humped-shaped term structure has also been observed. Next, the results of Black (1975) and Rubinstein (1985) are presented, which represent the initial empirical findings regarding the systematic pricing biases of the BS model across maturities.

Empirical Regularities of the Volatility Term Structure

Black (1975) reports that the BS model tends to overprice options with less than three months to maturity. He suggests several possible explanations for this pricing bias, but concludes that the observed price differences between market prices and BS

[74]The expression *volatility term structure* refers to the functional relationship and its plot.
[75]See Äijö (2008), p. 291.
[76]As mentioned in Chapter 2.1, this interpretation requires that the option pricing models hold.
[77]See Hafner (2004), pp. 39-40.

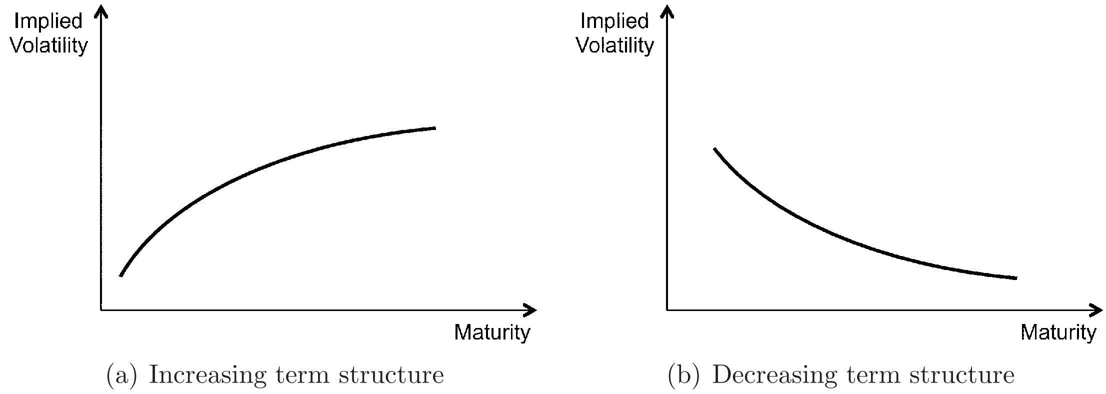


Figure 2.3.: Basic volatility term structure
Source: own illustration.

values cannot be fully explained. Thus, he remarks that Black and Scholes (1973) may have excluded something from the formula.[78] Rubinstein (1985) obtains strong evidence that the implied volatility of OTM calls increases if maturity declines. He finds mixed results for ATM calls. At the beginning of the sample period, the implied volatility of ATM calls increases for rising maturities, while at the end the opposite is observed.[79] Further evidence regarding the existence of volatility term structures in various options markets is provided by Franks and Schwartz (1991), Derman and Kani (1994a), Campa and Chang (1995), Backus et al. (2004), and Zhu and Avellaneda (1997), among others.

In addition to the existence of the volatility term structure, numerous studies have examined its time series behaviour. Notably, the response of the volatility term structure to current volatility changes has received some attention. Poterba and Summers (1986) analyse S&P 500 index options and report that long-term implied volatilities do not substantially react to volatility shocks. In contrast, Stein (1989), who examines whether the term structure is consistent with rational expectations, finds that the implied volatility of long-term options overreacts to short-term implied volatility changes.[80] He considers daily closing prices of ATM S&P 100 index options

[78]See Black (1975), p. 64.
[79]See Rubinstein (1985), p. 474.
[80]The analytical approach adopted by Stein (1989) is based on the notion that implied volatility should equal the average expected volatility over the remaining life-time of the option, if volatility

for the period from December 1983 to September 1987 and assumes that volatility follows a mean-reverting AR(1) process.[81] In summary, he concludes that investors overemphasise the effect of recent short-term implied volatility changes on long-term implied volatilities and neglect the low persistence of volatility shocks.[82]

Diz and Finucane (1993) reject the results of Stein (1989) by applying an alternative approach. As the residuals of the volatility model estimated by Stein (1989) are autocorrelated, they argue that the volatility model is misspecified. Using a sample that overlaps the time period investigated by Stein (1989), they report, contrary to Stein (1989), evidence for market underreaction. Additionally, the analysis of the non-overlapping sample provides no evidence for market under- or overreaction.[83] Furthermore, Heynen et al. (1994) question Stein's (1989) findings. Similar to Diz and Finucane (1993), they argue that the rejection of the hypothesis that implied volatility reflects rational expectations regarding future average volatility is due to the misspecification of the volatility process. They demonstrate that the hypothesis is not rejected if an EGARCH model is used to describe the dynamics of the underlying stock price. Overall, they suggest that the term structure is primarily determined by the level of unconditional volatility, which in their study is driven by an asymmetric GARCH process.[84]

While Stein (1989) and Heynen et al. (1994) perform joint tests of the expectation hypothesis and the option pricing model, Jiang and Tian (2010) use a model-free approach that does not require the specification of a particular option pricing model. They consider S&P 500 index options from June 1988 to December 2007. Their results call the previous findings of Stein (1989) and Diz and Finucane (1993) into

---

is unknown and options are non-redundant securities. Thus, his findings also require that the BS model hold.

[81] Stein (1989) performs so-called *term structure* tests which can be motivated by the following example. Assume that volatility oscillates around its mean (e.g. 20%) due to a strongly mean-reverting process. If the implied volatility of an option that expires in month is currently 30%, then the implied volatility of a three-month option should be on average lower than 30%, due to the mean-reversion effect. See Stein (1989), p. 1012.

[82] An alternative test that does not rely on the specification of a stochastic process for volatility confirms these results. See ibid., p. 1021.

[83] See Diz and Finucane (1993), p. 312

[84] See Heynen et al. (1994), pp. 50-51.

question, as they find no evidence for market misreaction. In particular, they demonstrate that the long-horizon overreaction in options markets documented by Stein (1989) and Poteshman (2000) is due to model misspecification.[85]

Stein (1989), Diz and Finucane (1993), and Heynen et al. (1994) analyse the volatility term structure based on options with two different maturities (in simplified terms, short- and long-term options). The approach proposed by Xu and Taylor (1994) allows the analysis of all traded options with arbitrary maturities.[86] They specify a two-factor model for the volatility term structure that takes short-term and long-term volatility into account. Their sample comprises daily data for four spot currency options that had been traded on the Philadelphia Stock Exchange from January 1985 to November 1989.[87] They report that volatility expectations move from short-term levels towards their long-term levels. By demonstrating that volatility expectations exhibit a half-life period of approximately four weeks, they provide evidence that the market does not expect volatility shocks to persist over longer horizons. Further, they find significant differences between short-term and long-term volatility expectations and hence evidence for the existence of a volatility term structure. Their findings also suggest that long-term expectations are time-varying. However, they note that short-term expectations change more rapidly. Moreover, they mention that the slope of the term structure changes frequently. In summary, they confirm the hypothesis that volatility expectations are formed rationally.[88]

Whereas the papers of Stein (1989), Heynen et al. (1994), and Xu and Taylor (1994) analyse the empirical regularities of the term structure for a given volatility model, Das and Sundaram (1999) use the reverse approach. They investigate the implications of different asset price processes for the form of the term structure. In particular, they employ jump-diffusion and stochastic volatility models. Their main result is that although each model class captures several empirical aspects of implied

[85] See Jiang and Tian (2010), p. 2359.
[86] However, as they estimate the term structure for nearest-the-money options, they do not consider the complete IVS.
[87] See Xu and Taylor (1994), p. 58.
[88] See ibid., p. 73.

volatilities quite well, both model classes fail to address all (important) empirical patterns in the data. Overall, they report that stochastic volatility models are better suited to replicating the typical shapes of the term structure of volatility smiles than jump-diffusions.[89]

In summary, Alexander (2008) notes that an upward sloping term structure is typically observed during calm market periods, while a downward sloping curve is generally found in volatile periods.[90] For instance, Äijö (2008) reports an extremely downward sloping term structure for the DAX, SMI, and Euro Stoxx 50 options markets after the September 11th attack. He highlights that the change in the shape of the term structure is due to short-term implied volatilities that increased dramatically. In contrast, long-term implied volatilities remained relatively stable.[91] Furthermore, the above-presented studies demonstrate that the volatility term structure varies over time. In the following, the empirical findings for the DAX options market are reviewed.

The DAX Volatility Term Structure

The above-cited study by Trautmann (1990) documents that the market prices of deep OTM call options with different maturities differ from corresponding BS values. This indicates a not-flat term structure for individual German stock options.[92] Further, Beinert and Trautmann (1995) demonstrate that the implied volatilities of individual stock options decrease if maturity rises. This relationship is more pronounced for ITM than for OTM options. For ATM call options, they report that the implied volatilities of short-term options are relatively low, increase for longer maturities, but decrease for options with the longest maturities.[93]

[89]See Das and Sundaram (1999), pp. 231-232.
[90]See Alexander (2008), p. 238.
[91]See Äijö (2008), p. 294.
[92]See Trautmann (1990), p. 99.
[93]See Beinert and Trautmann (1995), p. 14.

The results of Ripper and Günzel (1997) reveal that the implied volatilities of short-term deep ITM and OTM options (with maturities of up to ten days) are systematically higher than those for corresponding long-term options (with maturities from 18 to 95 days).[94] Fengler et al. (2002) analyse changes in the volatility term structure for DAX ATM options from March 1996 to December 1997. The DAX volatility term structure is derived from closing prices of eight VDAX subindices.[95] They find that the DAX term structure is downward sloping and changes over time during the sample period. In particular, they observe a temporary upward term structure shift and some clear term structural changes during the market downturn in the fall of 1997.[96]

Bolek (1999) considers the average DAX volatility term structure across all strike price classes, which is derived from DAX call options from the second half of 1995. He also provides evidence that DAX implied volatilities decrease if maturity rises. Similar to Fengler et al. (2002), he suggests that the DAX volatility term structure varies over time.[97] Wallmeier (2003) investigates transaction prices of DAX options and futures from 1995 to 2000 and reports contradictory results. He observes that the DAX volatility term structure for options with maturities from 30 to 120 days has a positive slope. Foresi and Wu (2005), who consider a longer sample period than Wallmeier (2003) ranging from June 1995 to May 2005, indicate that the implied volatilities of OTM put options increase at longer maturities. They observe the inverse pattern for OTM call options. The mixed results of Bolek (1999), Fengler et al. (2002), Wallmeier (2003), and Foresi and Wu (2005) are to some extent surprising, as their sample periods partially overlap. However, while all four studies investigate the DAX volatility term structure, their findings are based on partly different types of DAX options data. For instance, Bolek (1999) and Wallmeier (2003) use transactions data, Fengler et al. (2002) examine daily closing prices of volatility

[94]See Ripper and Günzel (1997), p. 475.

[95]The VDAX index family comprises volatility indices that reflect information on the implied volatility of DAX ATM options with different maturities. The implied volatilities used to calculate the VDAX indices are obtained from traded DAX options by inverting the BS formula.

[96]See Fengler et al. (2002), p. 18.

[97]See Bolek (1999), pp. 129-132.

indices, and Foresi and Wu (2005) analyse daily OTC volatility quotes. Additionally, Bolek (1999) investigates options with maturities of up to 12 months, Fengler et al. (2002) study subindices with maturities of 1, 2, 3, 6, 9, 12, 18 and 24 months, Wallmeier (2003) considers expiries of 30, 60, 90, and 120 days, and the sample in Foresi and Wu (2005) contains implied volatilities on options with maturities from 1 month to 5 years. However, the different shapes of the DAX volatility term structure demonstrate that it varies over time.

While the above studies are based on the BS model, Äijö (2008) examines the DAX volatility term structures based on the VDAX-NEW. The VDAX-NEW was introduced by the Deutsche Börse to make pure volatility tradable. While the VDAX is constructed from BS implied volatilities of DAX ATM options with a maturity of 45 days, the VDAX-NEW is not calculated using a particular option pricing model. The VDAX-NEW is derived from the market prices of traded DAX options. By this it can be replicated using a portfolio of DAX options. In addition to the VDAX, the VDAX-NEW not only considers the implied volatilities of ATM options, but also uses the implied volatilities of OTM options. Äijö (2008) reports on average a downward-sloping term structure for the period from January 2000 to December 2004. Further, the term structures exhibit considerable changes over time, as the slope ranges from negative to positive values during the sample period. Äijö (2008) also examines the relationship between term structure changes and market phases (see also the above result regarding the September 11th attack). He describes that at the end of the bear market in the second quarter of 2003, the slope changed (slowly) from a negative to a positive sign. Bearing these findings in mind, the contradictory results of Wallmeier (2003) highlighted above may be due to a long and extraordinary positive market upturn period during his observed sample period.[98]

[98]This possible explanation is qualified by the fact that the samples of Bolek (1999) and Fengler et al. (2002) also fell into periods of rising market prices, which however, were not as long as observed by Wallmeier (2003).

2.5.3. Dynamic Behaviour of Implied Volatilities

In the above Sections, it is often noted that the volatility smile and volatility term structure change over time. This Subsection picks up this finding and presents studies that address the time series behaviour of smiles and term structures. Similar to the organisation of prior passages, this Section first presents empirical studies on international options markets. The dynamic behaviour of DAX implied volatilities is described further below. The presentation of the selected international studies is organised as follows.

First, studies describing the dynamic features of individual implied volatility series with fixed moneyness and maturity are considered in chronological order. In particular, empirical results with respect to serial correlation, mean-reversion and the stationarity of implied volatilities are provided. Next, a study by Foresi and Wu (2005) analysing the time series behaviour of volatility smiles is presented. Thereafter, the results of Mixon (2007) and Äijö (2008), which examine the dynamics of the volatility term structure, are provided. In addition, the studies of Brooks and Oozeer (2002) and Christoffersen and Jacobs (2002) regarding volatility clustering of implied volatilities are briefly introduced. Finally, the findings of Gonçalves and Guidolin (2006), which investigate the dynamics of the complete IVS, are presented.

The Time Series Behaviour of the IVS in International Options Markets

Schmalensee and Trippi (1978) contributed an early study of time-varying implied volatilities and analyse weekly closing prices of six common stocks that had been traded on the CBOE from April 1974 to May 1975. Their time series analysis uses equally weighted-averages of implied volatilities, which compress the IVS to a single number. To describe the changes in the average implied volatilities, they calculate *corrected* first-order autocorrelation coefficients[99] and report (weak) evidence for

[99] They account for measurement errors induced by transaction costs and rounding errors. See Schmalensee and Trippi (1978), pp. 134-136.

negative serial correlation.[100] However, the validity of their study is limited by the fact that the performance of the BS option pricing model depends on moneyness and maturity.[101,102] Poterba and Summers (1986) avoid problems that are induced by determining weights. They use the CBOE Call Option Index for the period from January 1976 to June 1984.[103] By estimating weekly sample autocorrelations and sample partial autocorrelations for different lags, they suggest an AR(1) model to describe the (standardised) implied volatility series.[104] Based on the Dicky-Fuller test, they find that the implied volatility series is stationary.[105]

Harvey and Whaley (1991) study the time series properties of implied volatilities derived from transaction prices of short-term ATM S&P 100 index options. Their sample comprises the period from August 1988 to July 1989.[106] In particular, they investigate the effect of nonsimultaneous prices, the bid/ask spread, and infrequent trading on the time series behaviour of implied volatilities. They demonstrate that nonsimultaneous prices induce negative first-order autocorrelation in implied volatility changes. However, negative first-order sample autocorrelation is not completely eliminated after controlling for this effect. This problem is relevant for numerous studies using closing prices.[107] Furthermore, they show that the bid/ask price effect

[100] See Schmalensee and Trippi (1978), p. 135.

[101] The calculation of average weighted implied volatilities is based on equally weighted option prices.

[102] See Mayhew (1995), p. 9.

[103] The CBOE Call Option Index is calculated based on standardised stock option prices. For each CBOE stock, the option price is estimated for a hypothetical six-month ATM option using actual market prices. The option prices are standardised by dividing them by the underlying stock price. The index implicitly assumes a particular option with fixed moneyness and maturity. In so doing, the index captures the information of one point of the IVS. See Poterba and Summers (1986), p. 1148.

[104] The estimated first-order autocorrelation coefficient of weekly implied volatilities is 0.97. This implies that one year after a shock occurred $0.965^{52} \approx 16\%$ of the initial shock will be expected to be present in the data.

[105] See Poterba and Summers (1986), p. 1149.

[106] They use short-term options with a minimum of 15 days to maturity.

[107] For example, if the stock market closes before the options market, negative serial correlation can be induced by new information that is immediately incorporated into options market prices, but enter stock market prices with a time-lag of one night. Under the assumption that the information represents *good* market news, the implied volatility from call options, which is calculated based on stock closing prices for the previous day, is higher than it should be. On the next day, positive information is incorporated into the stock price and implied volatilities drop to normal levels. See Harvey and Whaley (1991), pp. 1553-1555.

leads to higher negative first-order serial correlation.[108] Again, negative serial correlation is not fully removed after taking bid/ask prices into account. In contrast to bid/ask spreads and nonsimultaneous prices, infrequent trading does not influence the serial correlation of implied volatilities. Drawing on these results, they argue that volatility levels should be stationary and mean-reverting.[109] As numerous early studies use closing prices, these studies are subject to these effects and their findings must be interpreted with caution.

In a subsequent study, Harvey and Whaley (1992) examine a larger sample of S&P 100 index options data that cover the period from October 1985 to July 1989. By calculating separate summary statistics for put and call options, they report that the implied volatility levels of both series are positively autocorrelated, which suggests persistence in the volatility level. The sample autocorrelations drop to zero at higher lags fairly quickly. This might indicates that the time series is stationary. Taking the differences of the implied volatilities, they find significantly negative autocorrelations at lags 1 and 2, which provides evidence against the hypothesis that volatility changes are unpredictable.[110] They use a linear model to predict volatility changes where implied volatility changes are regressed on certain informational variables. They observe that the coefficients of lagged implied volatility changes are significantly negative.[111] In combination with the observed first-order sample autocorrelation, the regression results provide evidence that the implied volatility level follows a mean-reverting process. Interestingly, they find that the changes in the implied volatility of call options are more predictable than those of put options. In summary, while their findings reveal that implied volatility changes are predictable

[108] Roll (1984) explains how the bid/ask spread can cause negative serial correlation in returns. See Roll (1984), pp. 1128-1130.

[109] See Harvey and Whaley (1991), p. 1558.

[110] See Harvey and Whaley (1992), p. 56.

[111] They estimate a regression model for the full period and for a second sample where the turbulent period from October 16th to October 30th, 1987 is excluded. The regression results show that the October 1987 market crash has a strong influence on the estimated coefficients. Therefore, Harvey and Whaley (1992) present an interpretation for the sample that excluded the crash.

in a statistical sense, an economic analysis proposes that arbitrage profits disappear when transaction costs are taken into account.[112]

Fleming et al. (1995) investigate changes in the CBOE Volatility Index (VIX) from January 1986 to December 1992.[113,114] They report that the changes in the average daily implied volatility are relatively low during the sample period. The sample autocorrelation structure of the daily volatility index changes are calculated for each year of the sample and vary substantially in the sample. In particular, they find positive and negative values for the first autocorrelation coefficient. For the complete sample period, they report a significantly negative first-order autocorrelation coefficient. Furthermore, when considering weekly implied volatility changes, they find a higher negative first-order autocorrelation coefficient, which indicates mean reversion behaviour in the index. Similar to Poterba and Summers (1986), they report that the first-order autocorrelation coefficient of the VIX series is approximately 97%.

The time series behaviour of volatility smiles is studied in the above-mentioned study by Foresi and Wu (2005), who consider daily OTC option quotes for 8 different maturities on 12 major equity indices. To examine the dynamic pattern of volatility smiles, for each index they estimate for each maturity and each day a second-order polynomial function of the form

$$\sigma(M;t,T) = \beta_0 + \beta_1 M + \beta_2 M^2 + \varepsilon \tag{2.18}$$

where $\sigma(M;t,T)$ represents the implied volatility at time t and for maturity T as a function of moneyness M and ε denotes the error term.[115] The parameter β_0 covers the level, β_1 the slope, and β_2 the curvature of the volatility smile. The daily estima-

[112] See Harvey and Whaley (1992), pp. 58-71.
[113] They use VIX changes, as VIX levels seem to follow a near-random walk.
[114] The VIX was constructed from the implied volatilities of eight OEX S&P 100 index options and had a constant 30-calendar day time to maturity. Note that in 2003, the CBOE changed the underlying index to the S&P 500 Index.
[115] Foresi and Wu (2005) define moneyness M as the ratio of the strike price K to the index level S.

tion of (2.18) yields a series of coefficients $(\beta_0, \beta_1, \beta_2)$ that is analysed with respect to its serial correlation structure. The estimated first-order sample autocorrelations of β_0 provide evidence that the implied volatility levels of the considered equity indices are persistent. Moreover, they find that the implied volatilities of short-term options (with maturities up to half a year) are less persistent than the volatilities of long-term options (with maturities up to 5 years). Further, the standard deviation of β_0 decreases if maturity rises. This demonstrates that the implied volatility level of short-term options varies to a greater extent over time than the level of long-term options. They report that all estimated slope coefficients are negative, which indicates a downward sloping volatility smile for all indices. In contrast to the time-variability of the volatility level, the standard deviation of β_1 increases for options with longer maturities. Thus, the slope of the volatility smile for long-term options shows more pronounced changes over time than the slope for short-term options. Additionally, with one exception, they find small values of β_2, implying that the shape of the volatility smile is similar to a straight line. Considering the standard deviation of β_2, the curvature of the volatility smile for long-term options moves to a greater extent than that for short-term options.[116]

Mixon (2007) investigates the dynamic behaviour of the volatility term structure for different equity indices. He analyses OTC data of ATM call options for the S&P 500, FTSE, DAX 30, CAC 40, and Nikkei 225 from May 1994 to October 2001. He finds that the slope of the average term structure is positive for 4 of the 5 indices.[117,118] However, he demonstrates that the slope of the term structure of the S&P 500 index changes several times during the sample period. For instance, the term structure was generally upward sloping during the period of low volatility levels from 1994 to 1996. However, with the increase in volatility levels at the end of 1997, the slope changed its sign and the term structure became downward sloping.[119] Therefore,

[116]See Foresi and Wu (2005), pp. 15-17.
[117]The exception is the Nikkei 225.
[118]The slope of the term structure is measured by the difference between the implied volatility of options with 1 and 12 months to maturity.
[119]Moreover, the slope changed from positive to negative values during the turbulent market period in Autumn 1998. It remained positive for most of the period from 1999 to 2000. After the

these findings provide further evidence that the slope of the term structure tends to be positive (negative) during phases of low (high) volatility.[120] Similar to Foresi and Wu (2005), he reports that the implied volatility level of short-term options (one month to maturity) varies to a greater extent over time than options with longer maturities (12 months to maturity).[121]

Äijö (2008) also presents evidence regarding the time series pattern of the volatility term structure for several popular equity indices. He considers daily index levels of the VDAX, VSMI and VSTOXX from January 2000 to December 2004. He reports that the sample autocorrelation coefficients are significantly positive up to five lags, which suggests mean-reverting behaviour in the implied volatilities.[122] To compare the persistence of a shock across the indices, he proposes the persistence measure $\ln 0.5/ln(\rho_1)$, where ρ_1 denotes the sample autocorrelation coefficient of lag one. Based on this measure, he finds that the half-life of a shock is between 31 and 38 days. Further, the results of the Dickey-Fuller and the Phillips-Perron tests confirm the stationary assumption for implied volatility levels and differenced implied volatilities.[123]

Gonçalves and Guidolin (2006) suggest a two-stage approach that makes it possible to model the dynamics of the complete IVS. They use a sample of daily closing prices for options on the S&P 500 index traded on the CBOE from January 1992 to June 1996. In a first step, similar to Dumas et al. (1998) they estimate daily cross-sectional models in which log implied volatilities are regressed on moneyness, squared moneyness, maturity, and an interaction term of moneyness and maturity. This regression yields a series of coefficients. In the second step, they fit a Vector Autoregressive (VAR) model to the multivariate time series of coefficients. With respect to the cross-sectional model, they find that the models typically provide an excellent

September 11 attacks the shape of the term structure was again downward sloping. See Mixon (2007), p. 340.

[120] See also the above-noted findings of Alexander (2008).

[121] See Mixon (2007), p. 340.

[122] See Äijö (2008), pp. 292.

[123] See ibid., pp. 293-295.

fit to the daily IVS.[124] However, the explanatory power of the cross-sectional models changed considerably over time. In contrast, the proposed cross-sectional model is able to reproduce the different observed IVS shapes. In accordance with Dumas et al. (1998), they demonstrate that the daily coefficients of the cross-sectional regressions are highly unstable over time. The results of the Ljung-Box (LB) test applied to the coefficient series confirm this finding. Furthermore, cross-correlograms indicate strong contemporaneous and lead and lag relationships among the estimated coefficients. For this reason, they fit a VAR model to the time series of coefficients.[125] Overall, they find that the two-stage model captures these relationships and the observed static and dynamic patterns of the IVS quite well. Although the two-stage model provides accurate forecasts from a statistical perspective, its economic performance is mixed.[126] In summary, these results suggest the existence of a regular dynamic structure of the IVS.[127] Having described the time series behaviour of the volatility smile, the term structure, and the IVS, the volatility clustering property of implied volatilities is described briefly in the following.

Volatility clustering property is well documented for numerous financial time series. While most studies consider the volatility clustering behaviour of stock returns, few articles address the volatility pattern of implied volatilities. One such study is Brooks and Oozeer (2002). They analyse the implied volatilities of options on Long Gilt Futures that were traded on the London International Financial Futures Exchange (LIFFE) from March 1986 to March 1996. They report that the implied volatility levels and the differenced series exhibit ARCH-effects. For this reason, they recommend the White-estimator to estimate the IVS based on a linear regression model. Another empirical study providing evidence for volatility clustering of implied volatility levels is Christoffersen and Jacobs (2002). They study weekly data on S&P 500 call options from June 1988 to May 1992. Their results also indicate that the volatility of implied volatility tends to cluster. Therefore, both studies indi-

[124] They report that the average adjusted R^2 is approximately 81%.
[125] See Gonçalves and Guidolin (2006), pp. 1598-1608.
[126] The simulation results concerning the economic performance of the approach are highly dependent on the assumptions regarding transaction costs and the definition of trading rules.
[127] See Gonçalves and Guidolin (2006), pp. 1628-1631.

cate that the volatility clustering effect also exists for implied volatilities. However, future research is necessary to support this hypothesis for a broader range of options markets.

Dynamic Pattern of DAX Implied Volatilities

This Section presents empirical findings on the dynamics of DAX implied volatilities. In particular, the study results of Hafner and Wallmeier (2000), Fengler et al. (2002), Wallmeier (2003), and Fengler (2012) are discussed.

Hafner and Wallmeier (2000) examine the time series pattern of DAX volatility smiles based on the transaction prices of DAX options. The DAX options were traded on the DTB/Eurex from January 1995 to October 1999. They estimate a spline regression model of the form

$$\sigma(M, D; t, T) = \beta_0 + \beta_1 M + \beta_2 M^2 + D\gamma_2(1 - 2M + M^2) + \varepsilon \tag{2.19}$$

which is conditional on a fixed maturity for each day. The dummy variable D takes the value one for $M > 1$ and zero otherwise.[128] It controls for the asymmetric strike pattern of implied volatilities. The regression is run for every day and for the two maturities.[129] They find that the estimated slope parameter $\hat{\beta}_1$ is negative on average, the curvature coefficient $\hat{\beta}_2$ is positive, and the asymmetry parameter $\hat{\gamma}_2$ is positive. Based on these findings, they conclude that the shape of the DAX volatility smile more closely resembles a skew during most of the sample period. Moreover, the regression results demonstrate that the volatility smile function is steeper and more convex for options with shorter maturities than for options with longer maturities.[130] This implies a more pronounced skew for short-term options. Further, the standard deviations of the estimated parameters exhibit substantial

[128] Here moneyness is defined as $M = \frac{K}{F(t,T)}$ where $F(t, T)$ denotes the DAX futures price at time t with maturity at time T.

[129] See Hafner and Wallmeier (2000), p. 14.

[130] Here, short-term (long-term) options are characterised by maturities which are lower equal (higher) than 45 days.

time-variation. According to Hafner and Wallmeier (2000), this does not indicate that the smile pattern changes distinctly during the sample period, as nearly identical smile patterns for ATM options can be generated by different parameters of the above smile function.[131] Furthermore, they report that the estimated parameters are highly correlated. Therefore, they do not investigate the dynamics of the regression coefficients as other studies do, but rather calculate two slope measures to analyse the time series behaviour of the DAX implied volatility smile.[132]

They consider the slope of the volatility smile for options from two moneyness regions.[133] First, both slope measures are positive over the complete sample period. This indicates that the shape of the DAX volatility smile more closely resembles a skew than a symmetric smile. Second, the upward movement of the slopes during the sample period implies that the DAX volatility skew became steeper. Third, as the values of the slope measure for options with a moneyness equal to 0.95 generally lie above the second slope measure (for options with moneyness equal to 1.05), the implied volatility skew is steeper for options within the moneyness boundaries of 0.95 and 1 than in the region from 1.0 to 1.05. In addition, they find that the sum of the two slope measures (called *total span*), which captures information regarding the smile profile, follows a stationary AR(1) process. As the estimated autocorrelation coefficient is near 0.98 for the complete sample period, they conclude that shocks to the volatility smile persist and die out slowly.[134]

In addition to the dynamic analysis of the slope of the volatility smile, Wallmeier (2003) measures and analyses the curvature of the volatility smile. He suggests a curvature measure based on the above two slope measures.[135] By calculating the

[131] For an example, see Hafner and Wallmeier (2000), p. 19.

[132] See ibid., pp. 19-20.

[133] The slope of the smile is calculated based on the difference between the implied volatility of a DAX option with a moneyness equal to 0.95 (or 1.05) and the implied volatility of a DAX ATM option. The slope measures reflect the average slope of the smile in the corresponding moneyness region.

[134] See Hafner and Wallmeier (2000), p. 32.

[135] The curvature measure is defined as the difference between the above two slope measures divided by the sum of both slope measures. See Wallmeier (2003), p. 190.

correlations between the changes in β_0[136] from (2.19), the relative total span[137], and the curvature measure, he finds that increasing ATM implied volatilities are accompanied by a decline in the relative total span and a reduction in the curvature. Therefore, the dynamics of the volatility smile are not fully captured by a proportional drift in the implied volatility level of ATM options.[138,139]

To identify common factors that govern the dynamics of implied volatilities, the above-mentioned study by Fengler et al. (2002) investigates 8 VDAX subindices from March 1996 to December 1997. The VDAX subindices are based on ATM DAX options with different maturities. Their analysis is based on differenced series of the VDAX subindices, as the nonstationarity hypothesis is only rejected for the differenced series and not for the VDAX subindex levels.[140] They report that the level and slope of the volatility term structure changed considerably during the market turmoil in 1997. Furthermore, using principal component analysis, they show that the total variation in the term structure of ATM DAX options can be primarily attributed to two risk factors.[141]

Fengler (2012) employs an alternative measure (called *skew measure*) for the slope of the volatility smile to that used by Hafner and Wallmeier (2000).[142] He applies the skew measure to DAX 1M and 1Y index option data from 2000 to 2008. He also identifies a volatility skew for the DAX options considered, as the skew measure is negative during the whole sample period. In particular, he reports that the skew measure increases during turbulent market periods for 1M options, which implies

[136] The intercept represents the general level of ATM implied volatility.

[137] They divide the total span by β_0 which reflects the absolute level of ATM implied volatility.

[138] See Wallmeier (2003), p. 193.

[139] Further, Wallmeier (2003) analyses the autocorrelation structure of β_0 and the total span. He finds that both variables can be modelled by an AR(1) process. As the curvature measure exhibits positive partial autocorrelations up to lag 4, he refrains from specifying an AR(1) model for the curvature. See Wallmeier (2003), p. 195.

[140] According to Fengler (2012), the data generating process for implied volatilities was often found to be nearly integrated. This characteristic makes it difficult to verify mean reversion in a statistical sense. See Fengler (2012), p. 123.

[141] See Fengler et al. (2002), p. 19.

[142] The measure is defined as $\frac{\partial \hat{\sigma}^2}{\partial M}\big|_{M=0}$ where M is given by $M = \log(K/S)$.

a steeper volatility smile.[143] Furthermore, he observes that the level of and the variation in the volatility smile of 1Y options are generally lower than in the case of the smile of 1M options. The co-movement of the DAX implied volatility level, the skew measure, and the term structure demonstrates that shocks across the IVS are highly correlated. Therefore, he concludes that the dynamics of the IVS can be adequately described by a small number of factors.[144]

Overall, the studies show that the volatility smile, the volatility term structure, and the IVS exhibits a systematic dynamic pattern that should be taken into account when predicting implied volatility. In the following, some potential explanations for the implied volatility patterns are presented.

2.6. Potential Explanations for the Stylised Facts of Implied Volatility

The literature suggests a series of explanations for the stylised facts of implied volatilities across moneyness and maturity presented above. As the existence of volatility smiles and term structures demonstrates that the assumptions of the BS model are violated, the explanatory approaches build on relaxing the BS assumptions. On the one hand, studies suggest that market microstructure effects, such as transaction costs, liquidity constraints, and information asymmetries, induce differences between observed market option prices and their corresponding BS values. On the other hand, a recent body of research proposes alternative stochastic processes for the underlying asset price, which differ from the geometric Brownian motion assumed by the BS model, to explain the observed volatility patterns, e.g., stochastic volatility or Lévy-processes.[145] First, this Section presents the explanation that attributes the volatility smile to stochastic volatility.

[143] The sample comprises the following volatile market periods: the September 11, 2001 attacks, the crash and aftermath of the dot-com bubble from 2001 to 2003, and the financial crisis 2008.
[144] See Fengler (2012), p. 120-123.
[145] See Hafner (2004), p. 56.

2.6.1. Stochastic Volatility

Stochastic volatility option pricing models (or in brief: *stochastic volatility models*) allow volatility to evolve stochastically over time via the introduction of an additional stochastic process. By relaxing the constant volatility assumption of the BS model, stochastic volatility models provide a flexible approach to capture the volatility smile phenomenon.[146] The first formal proof that a stochastic volatility model (the Hull and White (1987) option pricing model) produces a symmetric smile is given in Renault and Touzi (1996). They show that the volatility smile is a natural consequence of stochastic volatility. In particular, the resulting volatility smile is U-shaped and reaches its minimum for ATM options.[147] Hull (2006) briefly describes the relationship between stochastic volatility and volatility smiles as follows: stochastic volatility induces heavier tails in the *implied distribution* of asset prices than the corresponding lognormal distribution with the same mean and standard deviation.[148] First, a deep-out-of-the-money call option with strike price $K^{call}_{deep-OTM}$ is considered. Because of the heavier tails of the implied distribution, the probability that the asset price exceeds $K^{call}_{deep-OTM}$ is greater than for the lognormal distribution. Thus, the option price is higher given a leptokurtic distribution, which equates to higher implied volatility. The same argumentation can be employed to explain the higher price or, equivalently, the higher implied volatility of an OTM put option.[149] Thus, the implied volatility derived from model option prices based on a stochastic volatility model will exhibit a smile.[150] In addition to stochastic volatility models, GARCH-type option pricing models, e.g., Duan (1999), can also produce volatility smiles.[151]

[146] Alternatively, *local volatility models* relax the constant volatility assumption by modelling (local) volatility as a *deterministic* function of the asset price and/or time. See for instance Derman and Kani (1994a), Derman and Kani (1994b), Dupire (1994), and Rubinstein (1994).
[147] See Renault and Touzi (1996), p. 280.
[148] The implied distribution refers to the risk-neutral probability distribution of an asset price at future time t. It is determined by the volatility smile of options expiring at t.
[149] See Hull (2006), p. 378.
[150] See Alexander (2008), p. 271.
[151] See Duan (1999), p. 15.

While asset prices and volatility are by assumption uncorrelated in the Hull and White (1987) model, the Heston (1993) model allows for positive or negative correlation between the two processes. The non-zero price-volatility correlation influences the return distribution and produces a non-symmetric volatility skew. If the correlation is negative, the return distribution is left tailed and the volatility skew is negative. Alternatively, a positive price-volatility correlation generates a positive skew.[152] The magnitude of the smile is determined by the other parameters of the stochastic volatility model. In particular, the volatility of the volatility parameter affects the curvature of the smile.[153,154] However, Das and Sundaram (1999) demonstrate that stochastic volatility models require unrealistically high parameter values to produce the pronounced volatility smiles that are often observed for short-term options.[155] This theoretical finding is supported by an empirical study by Bates (1996b), who argues that jump-diffusion models are better suited to fit volatility smiles than stochastic volatility models. Thus, the class of jump-diffusion models and their appropriateness for capturing implied volatility patterns is described in the next Section.

2.6.2. Jumps

In addition to stochastic volatility, heavy-tailed return distributions can also be explained by jumps. Jumps interrupt the continuous asset price process and make it impossible to replicate the option pay off using a complete, riskless hedge portfolio. In this case, the principle of preference-free option valuation is no longer applicable.[156] As the shape of the volatility smile in the U.S. market has more closely resembled a skew or a sneer since the 1987 crash, Andersen and Brotherton-Ratcliffe

[152] See Alexander (2008), pp. 271-272 for a detailed description.
[153] See Hafner (2004), p. 57.
[154] Backus et al. (2004) argue that the increasing volatility term structure of ATM options can be explained by skewness and excess kurtosis. Their approach suggests that under certain assumptions, excess kurtosis pushes the implied volatility of ATM options downwards. Given increasing maturity, this effect declines and the implied volatility of ATM options increases. See Backus et al. (2004), p. 17.
[155] See Das and Sundaram (1999), p. 213.
[156] See Merton (1976), p. 132.

(2000) and others argue that this event caused market participants to increase their expected likelihood of a large downward movement.[157,158]

Merton (1976) proposed the first option pricing model to account for jumps. By combining a continuous price process with a discontinuous jump term, the stochastic differential equation for the asset price is given by

$$\frac{dS_t}{S_t} = (\mu - \lambda j_e)dt + \sigma dz_t + (J_t - 1)dq_t(\lambda) \tag{2.20}$$

where dz_t is a Wiener process and dq_t represents an independent Poisson process that generates the jumps. The parameter λ denotes the average number of jumps per unit time, J is an independent identically distributed random variable for the relative change in the asset price in the event of a jump, and $j_e = E(J-1)$ is the expected percentage change in the asset price. By imposing the assumption that the jump component of an asset's return is unsystematic, Merton (1976) implicitly assumes that jump risk can be diversified away, and thus the risk premium for jumps is zero.[159] This can be used to write the price of a European option determined by the Merton (1976) model as a weighted sum of BS prices.[160] Based on the jump-diffusion model and the assumption that $j_e = 0$, Merton (1976) shows that the pronounced volatility smiles of short-term options can be explained by jumps. In particular, he reports that the prices of deep-ITM and deep-OTM given by the Merton (1976) model exceed the BS value, while the Merton (1976) prices of ATM option are lower.[161] Furthermore, he finds that if investors expect a negative jump, the shape of the volatility smile becomes asymmetric.[162]

Das and Sundaram (1999) extend the results of Merton (1976) using the Bates (1996b) option pricing model in which jump and volatility risk are systematic and

[157] See Andersen and Brotherton-Ratcliffe (2000), p. 6.
[158] See also Bates (1991), pp. 1036-1037 and Pena et al. (1999), p. 1159.
[159] In this way, the Merton (1976) model represents an exception to the above rule that risk-neutral valuation is not possible in the presence of discontinuous jumps.
[160] See ibid., p. 135.
[161] See ibid., p. 140.
[162] See Wallmeier (2003), p. 62.

non-diversifiable. In particular, they show that jump-diffusion models are able to produce volatility smiles at short maturities under reasonable parameterisations. Unfortunately, jump-diffusions fail to generate realistic volatility smiles at long maturities, as the smile flattens out more rapidly than suggested by the empirical data. Finally, Das and Sundaram (1999) report that the term structure of ATM options from jump-diffusion models always exhibits an increasing shape, which contradicts the partly observed decreasing or humped term structure profiles. Therefore, jump-diffusion models make it possible to reproduce the skew pattern of short-term options but fail to explain skews at longer maturities or non-increasing volatility term structures. Thus, an option pricing model that combines stochastic volatility with jumps would provide sufficient flexibility to accurately match empirical skew and smile patterns.[163] However, combined models, e.g., Bates (1996b), contain more parameters than pure stochastic volatility or jump-diffusion models which negatively affects model parsimony.[164]

However, whether and to what extent jumps should be taken into account remains an ongoing question. Based on high-frequency data, Christensen et al. (2014) demonstrate that the effect of jumps on volatility is much lower than documented in the related literature. They argue that the relatively high attribution of jumps to asset price variability documented in the literature is often spurious, as these studies use low-frequency data. To my knowledge the consequences of this finding for option pricing have yet to be analysed. As the empirical Section of this thesis does not employ a jump-diffusion model, this argument is only mentioned here and will not be investigated further.

[163] See Das and Sundaram (1999), pp. 213-214.

[164] However, using S&P 500 index options data Bakshi et al. (1997) recommend option pricing models that account for stochastic volatility and jumps, because of their superior empirical performance and practicability.

2.6.3. Market Microstructure Effects

In addition to the consideration of alternative stochastic processes to explain volatility smiles and term structures, the literature also suggests that market microstructure effects influence the shape of implied volatilities across moneyness and maturity. First, the effect of market frictions is described. Then, the role of information asymmetries is examined.

Market Frictions

According to Whaley (2003), a frictionless market is characterised by the absence of trading costs and differential tax rates, unlimited borrowing and lending opportunities at the risk-free rate, no short selling restrictions, and the possibility to trade at any time and in any quantity.[165] If any of these market assumptions is violated, then the arbitrage mechanism ensuring that, in the BS world, the option price is equal to the price of the *hedging portfolio* is affected.[166] As a result, the value of the BS hedging portfolio is influenced and continuous, dynamic re-hedging becomes more complicated or even impossible.

The effect of transaction costs on option pricing is analysed in Gilster and Lee (1984), Leland (1985), Boyle and Vorst (1992), and Longstaff (2005), among others. Wallmeier (2003) notes that transaction costs can contribute to explaining volatility smiles if option traders only perform dynamic hedging strategies for certain options.[167] Studies investigating whether transaction costs are one of the determinants of the volatility smile provide mixed results. While Longstaff (1995) reports empirical evidence that pricing biases in S&P 100 index options can be induced by market frictions, Constantinides (1996) provides a theoretical analysis and argues that transaction costs cannot account for volatility smiles. Furthermore, Pena et al.

[165] See Whaley (2003), p. 1140.

[166] Black and Scholes (1973) assume that the payoff of an option contract can be replicated by a hedging portfolio with the same payoff. Based on a dynamic hedging argument, they are able to transform an option into a risk-free instrument.

[167] See Wallmeier (2003), p. 145.

(1999) consider Spanish IBEX-35 index options from January 1994 to April 1996 and find that transaction costs play a major role in explaining the curvature of the volatility smile.[168]

Market liquidity is an additional facet of market friction used in literature to explain volatility smiles. As option contracts with high implied volatilities are often less liquid, it is reasonable to conclude that the implied volatility patterns are caused by liquidity effects. As mentioned above, some authors suggest that the pronounced volatility smiles that have occurred since the 1987 stock crash are induced by excess demand for OTM puts. The excess demand for OTM puts is driven by (institutional) investors' need for portfolio insurance.[169] Because investors are faced with position limits and hedging costs, the supply of OTM puts is restricted, and as a result the excess demand cannot be satisfied. This leads to higher option prices and therefore implied volatilities.

This explanation is supported by the results of Bollen and Whaley (2004), who investigate the relationship between net buying pressure and the shape of the volatility smile.[170] They analyse S&P 500 index options data from June 1988 to December 2000 and confirm that implied volatility changes depend on the net buying pressure from public order flows. In particular, they report that the shape of the S&P 500 index smile can be related to the buying pressure for index puts. Although arbitrage profits should induce traders to sell OTM put options and replicate them synthetically, Isaenko (2007) argues that short-sale constraints on trading stocks and derivatives hamper this mechanism. Further, the results of Foresi and Wu (2005) support the portfolio insurance argument as an explanation for volatility smiles. They report that downside movements in major equity indices appear to be highly globally correlated due to worldwide market linkages. As a consequence, this global

[168]They measure transaction costs based on the bid-ask spread.
[169]See for instance Boyle and Vorst (1992), p. 285.
[170]They define net buying pressure as the difference between the number of buyer-motivated contracts and the number of seller-motivated contracts that are traded each day. A buyer-motivated (seller-motivated) trade is characterised by an execution price above (below) the prevailing bid/ask midpoint.

downside risk cannot be easily diversified. Thus, the price of OTM puts that protect against downside movements contains a corresponding risk premium.[171]

The relationship between volatility smiles and market liquidity is also addressed in Grossman and Zhou (1996), Platen and Schweizer (1998), and Frey and Patie (2002). In particular, these studies suggest that volatility smiles are related to feedback effects from dynamic hedging strategies. In an equilibrium framework, Grossman and Zhou (1996) assume that the demand for portfolio insurance is exogenously driven by a group of investors. In the context of their model, they explain why OTM put options exhibit higher implied volatilities than ITM puts. They conclude that volatility smiles reflect the equilibrium price impact of portfolio insurance. Platen and Schweizer (1998) develop a model in which the stock price incorporates the (demand) effect of hedging strategies into account. By using this model they provide numerical evidence that implied volatilities are due to feedback effects from hedging strategies. Frey and Patie (2002) criticise the approach of Platen and Schweizer (1998), as the latter employ an implausible model parameterisation to explain volatility smiles. However, in a related model, Frey and Patie (2002) propose that the volatility smile pattern is induced by market illiquidity. In particular, based on their model, they demonstrate that the lack of market liquidity due to a large market downturn leads to volatility skews. The assumption that large downward or upward asset price movements reduce the level of market liquidity is quite plausible.[172] In the next Section, the effect of information asymmetry on implied volatility is described.

Information Asymmetry

Prior research by Back (1993) and Nandi (2000) established a relationship between information quality and option prices.[173] Back (1993) develops an equilibrium model in which an informed agent can trade an option and the underlying asset, e.g.,

[171]See Foresi and Wu (2005), p. 11.
[172]See Fengler (2004), p. 45.
[173]See also Easley et al. (1998) and Shefrin (1999).

a stock. As option trades convey different information than stock trades in this model, issuing options affects the underlying stock price. Due to the change in the information flow, the volatility of the underlying assets becomes stochastic.[174] To ensure that trading takes place in both markets, he introduces so-called *noise* or *liquidity traders* that trade for non-informational reasons and thereby this provide liquidity to the market. In particular, he assumes that liquidity trades in the option and stock markets are not fully correlated. Based on Back's (1992) finding that the volatility pattern of the underlying asset depends on the structure of the liquidity trades, he suggests that the option price (or implied volatility) is influenced by the pattern of liquidity trades in the option.[175]

Nandi (2000) proposes a multiperiod model of asymmetric information that postulates a relationship between the level and curvature of the volatility smile and net options order flows. This model assumes that an agent has private information on the future volatility of the asset price but does not know the future level of the asset price. In particular, he finds that higher net options order flows lead to higher levels of implied volatility, which reflects an increasing pricing bias in the BS model. In addition, the options order flow also affects the curvature of the smile. The model's mechanics can be described as follows: because options order flows provide information on future volatility, increasing net options order flows reflect an increase in future volatility. Thus, the market maker who tracks options order flows to obtain information on future volatility posts higher option prices to avoid losses from trading with informed investors. This induces a greater degree of mispricing in the BS model, as the BS model does not take order flows into account.[176]

Buraschi and Jiltsov (2006) and Vanden (2008) extend the analysis of the relationship between information quality and option prices. They investigate how implied volatilities are related to changes in information quality. While Buraschi and Jiltsov (2006) consider the impact of public information, Vanden (2008) analyses how costly

[174]See Back (1993), pp. 450-454 for a detailed description.
[175]See ibid., pp. 435-438.
[176]See Nandi (2000), pp. 216-217.

private information can affect option prices. Buraschi and Jiltsov (2006) suggest that heterogeneous information can explain the volatility smile. They develop an equilibrium model in which they assume that agents are rational, have identical preferences and initial wealth, but incomplete and heterogeneous information. In the model, the agents are forced to form expectations regarding future dividends, as Buraschi and Jiltsov (2006) assume that the dividend growth rate is stochastic. The heterogeneous information assumption implies that the agents select different optimal portfolios due to their different expectations. In equilibrium, the agents who expected low dividend growth rates ask for portfolio insurance in form of OTM puts from the agents with higher dividend growth rate estimates. In contrast, agents with high dividend growth rate estimates demand OTM call options from the other agents. Because the agents are risk averse, their marginal utility is higher (lower) in bad (good) states of the economy. It follows that the cost of an OTM put exceeds the cost of an OTM call option, and thus a volatility smile occurs.[177]

Vanden (2008) proposes a multiperiod model in which agents can purchase private information regarding the underlying asset during multiple trading rounds. He shows that the change of information quality is the main driver of the risk-neutral distribution of asset returns. Therefore, variations in information quality and information acquisition costs over time affect the dynamics of implied volatilities. This provides an explanation for the puzzle wherein option prices occasionally change even though the price of the underlying asset remains constant.[178] Furthermore, he argues that shifts in information quality can also explain changes in the volatility term structure that have previously been attributed to market overreaction (e.g., Stein (1989)). Moreover, he suggests that investor overconfidence drives the shape and dynamics of the volatility term structure.[179]

[177] See Buraschi and Jiltsov (2006), pp. 2841-2846.
[178] See Vanden (2008), p. 2664.
[179] See ibid., p. 2637.

2.6.4. Conclusion

While the literature on potential explanatory approaches is enormous, no solution appears sufficient to fully explain the implied volatility patterns.[180] Bertsch (2008) notes that the explanation could differ across markets. For instance, relative to individual stocks, it is less likely that a stock index smile is induced by jumps, as a stock index represents an aggregation of individual stocks that generates a smoothing effect due to averaging.[181,182] Moreover, it is reasonable to imagine that a specific smile/skew pattern could be explained by a combination of several approaches. As it is difficult to separate the effects of different explanatory factors, this will present a challenge for future research.[183]

[180] See Jackwerth (2004), p. 8.
[181] See Saunders (1997), p. 66.
[182] Branger and Schlag (2004) investigate the effect of individual stock characteristics on the index smile. They demonstrate that the steepness of the volatility smile can be fully explained by the dependence structure of the stocks that constitute the index.
[183] See Fengler (2004), p. 44.

3. Analysis of DAX Implied Volatilities

Having presented the BS model and research findings regarding its pricing biases, the aim of this Chapter is to analyse mispricing in DAX options based on the BS model. Depending on the results, it will be clear whether and to what extent the BS model represents an accurate model of DAX option prices. If the BS model works well, this will provide an argument in favour of the use of BS implied volatilities as volatility forecasts. If the BS model performs poorly, the theoretical basis for the predictive ability of BS implied volatilities would be reduced to a heuristic rule.

To analyse BS implied volatilities across moneyness and maturity, it is necessary to construct a smooth IVS on a prespecified grid. Smoothing the IVS ensures that one only receives a single implied volatility for each given strike and maturity combination. Further, the effects of recording errors are reduced. A smooth IVS can be obtained by parametric or non-parametric approaches.[1] The basic concepts of the two approaches are presented in the next Section. Thereafter, the stylised facts of the DAX IVS that reflects the mispricing behaviour of the BS model are documented.

[1]See Fengler (2004), p. 97.

3.1. Methods for Smoothing the IVS

3.1.1. Introduction

Several smoothing methods have been developed for implied volatilities to derive the risk-neutral distribution from option prices. If the risk-neutral distribution of the underlying asset price is known, the price of any derivative written on the asset with the same time to maturity can be determined. Breeden and Litzenberger (1978) demonstrate that the risk-neutral distribution can be recovered from option prices.[2] Calculating the risk-neutral distribution based on Breeden and Litzenberger's (1978) result requires that call options with the same maturity and a continuum of strike prices from zero to infinity are available.[3] However, in practice, option contracts only exist at discretely spaced strike levels. Therefore, several inter- and extrapolation techniques have been suggested to complete the call price range (or equivalently implied volatilities).[4] The smoothing methods described below are taken from these fields of research.

Fengler (2004) categorises the smoothing techniques into parametric and non-parametric methods. Parametric approaches often employ polynomial specifications to fit the IVS. Some selected specifications, such as Shimko (1993) and Dumas et al. (1998) are presented in this Section. In addition, cubic splines, which provide greater flexibility, are introduced (e.g., Campa et al. (1997) and Hafner and Wallmeier (2000)). While parametric methods require the specification of a certain regression equation, non-parametric methods are based on locally averaging the data. Of the class of non-parametric methods, the Nadaraya-Watson estimator and higher order local polynomial smoothing, which are often applied to estimate the IVS, are presented below.[5]

---

[2] In particular, the second partial derivative of the call option pricing formula is used to derive the risk-neutral distribution.

[3] See Saunders (1997), p. 74.

[4] See Jackwerth (2004).

[5] In addition, Fengler (2004) proposes a semiparametric factor model to estimate the IVS that takes the degenerated string structure of the IVS data into account.

3.1.2. Parametric Methods

Regression models have been widely used to fit the implied volatility function. Numerous specifications have been suggested in the literature. This Section provides an overview of selected regression equations that haven been applied in the previous empirical literature.[6]

Shimko (1993) developed the first approach to smooth implied volatilities via a polynomial regression equation. He suggests a simple quadratic polynomial of the form

$$\sigma(K) = \alpha_0 + \alpha_1 K + \alpha_2 K^2 \tag{3.1}$$

to fit the implied volatilities.[7] Brunner and Hafner (2003) apply Shimko's (1993) method to DAX option contracts traded on the DTB/Eurex in the year 2000 and find that the method provides a good fit.[8] In an earlier study, Ripper and Günzel (1997) regress DAX implied volatilities on moneyness and squared moneyness for different maturities. Their sample comprises settlement prices for DAX options from 1995 to 1996. They report that the model parameters are significant and the coefficient of determination is high.[9]

Dumas et al. (1998) extend Shimko's (1993) approach by adding maturity and an interactive term to the quadratic polynomial. Dumas et al. (1998) examine different

[6]As some of the regression models were introduced in Section 2.5.3, they are only briefly addressed in this Section.

[7]A complete description of Shimko's (1993) method to extract the risk-neutral distribution from option prices or implied volatilities can be found in Walter (2008), p. 25.

[8]See Brunner and Hafner (2003), p. 95.

[9]See Ripper and Günzel (1997), p. 474.

structural forms of the volatility function and find that the following models deliver the best results[10]:

$$\text{Model 1}: \quad \sigma = \alpha_0 + \alpha_1 K + \alpha_2 K^2 \tag{3.2}$$

$$\text{Model 2}: \quad \sigma = \alpha_0 + \alpha_1 K + \alpha_2 K^2 + \alpha_3 T + \alpha_5 KT \tag{3.3}$$

$$\text{Model 3}: \quad \sigma = \alpha_0 + \alpha_1 K + \alpha_2 K^2 + \alpha_3 T + \alpha_4 T^2 + \alpha_5 KT. \tag{3.4}$$

Moreover, Tompkins (2001), Aït-Sahalia (2002), and Brunner and Hafner (2003) use third-order polynomials to estimate the implied volatility function. Brunner and Hafner (2003) compare the performance of Shimko's (1993) quadratic model to that of a cubic form for DAX options. They find that the cubic form is better suited to model DAX implied volatilities.[11] The following table presents, selected empirical studies using the above smoothing technique. The table also contains studies that employ spline methods, as spline-based methods have some desirable characteristics that are described below.

[10] These volatility functions belong to an option pricing model suggested by Dumas et al. (1998). They are determined by minimising the sum of squared errors between model option prices and actual, observed option prices.

[11] See Brunner and Hafner (2003), p. 95.

study	**volatility function**	**data**
Ncube (1996)	$\ln(\sigma) = \beta_0 + \beta_1 T + \beta_2 T^2 + \beta_3 K + \beta_4 D + \varepsilon$ where $D = 1$ for puts and $D = 0$ for calls	daily data, FTSE 100 index options, Nov 1989 to Mar 1990
Ripper and Günzel (1997)	$\sigma = \beta_0 + \beta_1 M + \beta_2 M^2 + \varepsilon$	daily data, DAX index options, Jan 1995 to Dec 1996
Pena et al. (1999)	$\sigma = \beta_0 + \varepsilon$ $\sigma = \beta_0 + \beta_1 M + \varepsilon$ $\sigma = \beta_0 + \beta_1 M + \beta_2 M^2 + \varepsilon$ $\sigma = \beta_0 + \beta_1 U + \beta_2 D^2 + \varepsilon$ $\sigma = \beta_0 + \beta_1 U + \beta_2 M^2 + \varepsilon$ $\hat{\sigma} = \beta_0 + \beta_1 U + \beta_2 M^2 + \beta_3 D + \varepsilon$ where $D = 0$ if $K < 1$ and $D = K$ if $K \geq 1$, $U = K$ if $K < 1$ and $U = 0$ if $K \geq 1$.	transaction data, IBEX35 index options, Jan 1994 to Apr 1996
Hafner and Wallmeier (2000)	$\sigma = \beta_0 + \beta_1 M + \beta_2 M^2 + D\gamma_2(1 - 2M + M^2) + \varepsilon$ where $D = 1$ for $M > 1$ and $D = 0$ for $M \leq 1$	transaction data, DAX index options, Jan 1995 to Dec 1999
Brunner and Hafner (2003)	$\sigma = \beta_0 + \beta_1 M + \beta_2 M^2 + D\beta_3 M^3 + \beta_4 \sqrt{T} + \beta_5 M\sqrt{T} + \varepsilon$ where $D = 1$ for $M > 0$ and $D = 0$ for $M \leq 0$	transaction data, DAX index options, Jan to Dec 2000
Christoffersen and Jacobs (2004)	$\sigma = \beta_0 + \beta_1 K + \beta_2 K^2 + \beta_3 T + \beta_4 T^2 + \beta_5 KT + \varepsilon$	daily data, S&P 500 index call options, Jun 1988 to May 1991
Hafner (2004)	$\sigma = \beta_0 + \beta_1 M + \beta_2 M^2 + \beta_3 \ln(1 + T) + \beta_4 M \ln(1 + T) + \beta_5 M^2 \ln(1 + T) + \varepsilon$	transaction data, DAX index options, Jan 1995 to Dec 2002
Gonçalves and Guidolin (2006)	$\ln(\sigma) = \beta_0 + \beta_1 M + \beta_2 M^2 + \beta_3 T + \beta_4 MT + \varepsilon$	daily data, S&P500 index options, Jan 1992 to Jun 1996
Berkowitz (2010)	$\sigma = \beta_0 + \beta_1 K + \beta_2 T + \beta_3 K^2 + \beta_4 T^2 + \beta_5 KT + \varepsilon$	daily data, S&P 100 index options, different samples

Note: Due to readability it is refrained from using the index IV in σ_{IV} for implied volatility.

Table 3.1.: Studies based on parametric volatility functions

Despite its use in several studies, Brunner and Hafner (2003) argue that a simple quadratic polynomial cannot adequately capture the structure of implied volatilities. However, Jackwerth (2004) notes that using higher order polynomials can introduce oscillatory effects. Therefore, he recommends employing *splines*, as they are able to produce smooth implied volatility functions and avoid oscillatory behaviour.[12] Additionally, Campa et al. (1997) argue that splines also provide greater flexibility to match the shape of the implied volatility function. This ability is described below.

[12]To produce smooth, risk-neutral probability distributions, splines need to be of an order higher than three. See Jackwerth (2004), p. 24.

Splines consist of polynomial segments that are allowed to change their shape over each segment. The polynomial segments are spliced at the knot points under certain conditions. For instance, cubic splines are based on cubic segments that are matched together at the knots under the condition that the function is twice differentiable at these points.[13] The cubic spline function for a given set of knot points $\{(x_i, \sigma_i)\}_{i=1}^{n}$ is

$$CS_j(x) = \sigma_j + \beta_j(x - x_j) + \gamma_j(x - x_j)^2 + \delta_j(x - x_j)^3 \tag{3.5}$$

where $CS_j(x)$ represents a smooth implied volatility function of x (e.g., a strike price) that is twice differentiable at the knots and defined for the subinterval $[x_j, x_{j+1}]$. *Natural cubic splines* are obtained under the boundary conditions

$$CS''(x_1) = \cdots = CS''(x_n) = 0 \tag{3.6}$$

and *clamped cubic splines* require

$$CS'(x_1) = f'(x_1), \ldots, CS'(x_n) = f'(x_n) \tag{3.7}$$

where $\sigma_i = f(x_i)$ for a given set of points x_i $(i = 1, ..., n)$.[14] Jackwerth (2004) highlights that a careful determination of the knots is necessary. While the use of additional knots improves the fit, an excessive number of knots will lead to overfitting and instable results.[15]

In the literature, splines have been applied by Bates (1991), Campa et al. (1997), Campa et al. (1998), Bliss and Panigirtzoglou (2002), Andersen and Wagener (2002), Hafner and Wallmeier (2000), Jiang and Tian (2005), and Fengler (2009), among others, to options on stock indices, stock index futures, exchange rates, and interest rate futures.[16] Bates (1991) fits constrained cubic splines to the ratio of the options price to the futures price as a function of the ratio of strike price to the futures

[13] See Andersen and Wagener (2002), p. 16.
[14] See Rouah and Vainberg (2007), pp. 28-31 for illustrations.
[15] See Jackwerth (2004), p. 24.
[16] To ensure the absence of arbitrage in the IVS, Fengler (2009) proposes an approach for arbitrage-free smoothing of implied volatilities.

price. By estimating separate cubic splines for calls and puts on S&P 500 futures for the period from 1985 to 1987, he reports an excellent fit before the 1987 crash, but enormous standard errors after the crash.[17] Whereas Bates (1991) applies cubic splines to the ratio of option prices and futures prices, Campa et al. (1997) fit splines to implied volatilities. Their sample contains OTC options data on eight European exchange rates from April 1996 to December 1996. They report that splines are able to produce asymmetrical smiles.[18] The contribution of Campa et al. (1997) lies in the modification of Shimko's (1993) method to derive the risk-neutral distribution, as they replace quadratic polynomials with cubic splines.

In a subsequent paper, Campa et al. (1998) apply a natural cubic spline to liquid OTC currency options on dollar-mark, dollar-yen, and several European cross rates from April 1996 to March 1997. They find that it successfully replicates the volatility smile.[19] In addition to the advantages of the approach, Campa et al. (1998) mention that the extrapolated implied volatilities outside the range of the observed strike prices can increase dramatically when splines are used.[20] Bliss and Panigirtzoglou (2002) also use a natural spline method but assume that implied volatility is a function of the options' deltas. They consider FTSE 100 index options and short sterling futures options during 1997. They find that natural splines provide an excellent fit to the data. Andersen and Wagener (2002) investigate Euribor futures option prices from March 2001 to June 2001. They propose higher order polynomials with sparse knot sets, as cubic splines can lead to non-differentiable risk-neutral distributions.[21] While the previous studies implement spline methods to derive the risk-neutral distribution from option prices (or implied volatilities), the primary objective of the following papers is to smooth implied volatilities. The smoothed implied volatilities

[17] See Bates (1991), p. 1020.

[18] The robustness of their results has been questioned, as they use five observations to derive the smile.

[19] See Campa et al. (1998), p. 131.

[20] See ibid., p. 128.

[21] Andersen and Wagener (2002) argue that this is an undesirable feature that should be avoided, as it is induced by the method selected to extract the risk-neutral distribution. See Andersen and Wagener (2002), p. 4.

are used to calculate model-free volatility (Jiang and Tian (2005)), local volatility Fengler (2009)), or to directly analyse the IVS (Hafner and Wallmeier (2000)).

Jiang and Tian (2005) is closely related to this study. To calculate model-free implied volatility, they assume that option prices from a continuous strike price range are available. As this is not realistic in practice, Jiang and Tian (2005) apply cubic splines to interpolate between the available strike prices of S&P 500 index options from June 1988 to December 1994. While the above studies on stock index options consider foreign stock markets, Hafner and Wallmeier (2000) and Fengler (2009) use DAX options data, which is of particular interest for this study.

Hafner and Wallmeier (2000) analyse DAX implied volatilities for the period from 1995 to 1999[22] and apply a spline function of the form

$$\sigma(M, D) = \beta_0 + \beta_1 M + \beta_2 M^2 + D(\gamma_0 + \gamma_1 M + \gamma_2 M^2) + \varepsilon \tag{3.8}$$

where D is a dummy variable with

$$D = \begin{cases} 0, \ M \leq 1 \\ 1, \ M > 1. \end{cases} \tag{3.9}$$

To control for heteroscedasticity, they estimate the regression equation using a weighted least squares estimator. To guarantee a continuous, differentiable, and smooth function, they impose certain additional restrictions.[23] They find that the above model is able to capture the asymmetric strike pattern of DAX implied volatilities. In addition, they report that the regression explains, on average, approximately 95% of the cross-sectional variation in implied volatilities.[24]

Similar to Bates (1991), Fengler (2009) suggests an approach that uses natural cubic splines to smooth call prices for a single time-to-maturity due to their computational

[22] They consider DAX options with a (hypothetical) remaining lifetime of 45 days.
[23] See Hafner and Wallmeier (2000), pp. 13-17.
[24] See ibid., pp. 31-33.

advantages. In particular, his approach ensures that the estimated IVS is arbitrage-free.[25] However, while parametric methods are used in several studies, Fengler (2004) argues that they seem to exhibit weaknesses when capturing the typical features of the IVS. Therefore, nonparametric methods have recently been suggested to overcome these shortcomings.[26] They are presented in the next Section.

3.1.3. Nonparametric Methods

Motivation

Economic theory often suggests the direction of influence between two related variables, but does not specify the functional form of the relationship. Parametric methods, such as the above described regression models, impose assumptions regarding the form of the functional relationship. In particular, they assume a specific model for the conditional expectation. In this manner, parametric models make it possible to extrapolate the data or to test restrictions that are postulated by theory. However, the advantages of parametric models come at the cost of specification errors, which induce inconsistent estimates. Further, hypothesis tests of the model are joint tests of the theory and the assumed functional form of the relationship. For this reason, nonparametric methods have been proposed to overcome these drawbacks.[27] Recently, nonparametric methods have been applied by Aït-Sahalia and Lo (1998), Rosenberg (1999), Cont and da Fonseca (2002), Fengler et al. (2003), Detlefsen (2007), Benko et al. (2009), and Birke and Pilz (2009).

Nonparametric methods make it possible to examine how the dependent variable reacts to changes in the independent variables without assuming a particular model for the conditional expectation. In the following, the basic concept of the standard Nadaraya-Watson kernel regression estimator is described, which has been

[25] He notes that arbitrage violations can lead to negative transition probabilities, which induce pricing biases.
[26] See Fengler (2004), p. 97.
[27] See Aït-Sahalia and Duarte (2003), p. 9.

used in numerous studies to estimate the IVS. Thereafter, a more general family, local polynomial kernel estimators, which nests the Nadaraya-Watson estimator, are presented.

The Nadaraya-Watson Estimator

To approximate the unknown functional relationship m between an explanatory variable X (e.g., moneyness and/or maturity) and a response variable Y (e.g., implied volatility), the model

$$Y = m(X) + \varepsilon \tag{3.10}$$

is assumed, where X is stochastic, the strict exogeneity assumption holds, and ε is i.i.d. with zero expectation and variance $\sigma^2(X)$. Therefore, the conditional expectation of Y given $X = x$ is

$$E[Y|X = x] = m(x) = \frac{\int y f(y,x) dy}{f_x(x)} \tag{3.11}$$

where $f(y, x)$ denotes the joint density of (Y, X) and $f_x(x)$ represents the marginal density of X. Thus, the unknown regression function can be estimated using the kernel density estimates of the joint and marginal densities. The resulting estimator was independently developed by Nadaraya (1964) and Watson (1964) and has the formula

$$\widehat{m}(x) = \frac{1}{n} \sum_{i=1}^{n} \frac{K_h(x - x_i)}{\frac{1}{n}\sum_{j=1}^{n} K_h(x - x_j)} y_i = \frac{1}{n} \sum_{i=1}^{n} w_{i,n}(x) y_i \tag{3.12}$$

where $K_h(\cdot)$ is a *kernel function* that weights the data.[28]

Equation (3.12) shows that kernel smoothing is based on the notion of locally averaging the data.[29] The response variable y_i is locally averaged with the weights

$$w_{i,n}(x) = \frac{K_h(x - x_i)}{\frac{1}{n}\sum_{j=1}^{n} K_h(x - x_j)}. \tag{3.13}$$

[28]See Schnellen (2007), pp. 10-11.
[29]See Jackwerth (2004), p. 22.

Different kernel functions have been proposed to implement different weighting schemes. Usually, the kernel functions are continuous, positive, bounded, symmetric and integrate to one. Typical kernel functions are the *Gaussian* kernel, which is given by

$$K_h(u) = \frac{1}{\sqrt{2\pi}} e^{-u^2/2}, \tag{3.14}$$

the *Epanechnikov* kernel

$$K_h(u) = \frac{3}{4}(1-u^2) \qquad \mathbf{1}(|u| \leq 1), \tag{3.15}$$

and the *quartic* kernel

$$K_h(u) = \frac{15}{16}(1-u^2)^2 \qquad \mathbf{1}(|u| \leq 1).^{30} \tag{3.16}$$

The estimation of the IVS requires a multidimensional kernel function. A multidimensional kernel can be constructed by taking the products of multiple univariate kernels

$$K_h(u_1, \ldots, u_d) = \prod_{j=1}^{d} K_h^{(j)}(u_j).^{31} \tag{3.17}$$

For instance, Aït-Sahalia and Lo (2000), Härdle et al. (2002), and Cont and da Fonseca (2002) use a two-dimensional Nadaraya-Watson estimator that takes the form

$$\hat{\sigma}_t(M,T) = \frac{\sum_{i=1}^{n} K_h(M-M_i, T-T_i)\sigma_t(M_i,T_i)}{\sum_{i=1}^{n} K_h(M-M_i, T-T_i)} \tag{3.18}$$

where

$$K_h(x,y) = (2\pi)^{-1} \exp(-x^2/2h_1) \exp(-y^2/2h_2) \tag{3.19}$$

denotes a Gaussian kernel with bandwidth parameters h_1 and h_2.[32] Alternatively, Fengler et al. (2003) apply a two-dimensional Nadaraya Watson estimator with a quartic kernel.

[30] See Fengler (2004), pp. 99-100.
[31] See ibid., p. 100.
[32] See Cont and da Fonseca (2002), p. 49.

While the coefficients of a correctly specified parametric model are unbiased, nonparametric estimators are biased. The bias is defined as $\text{Bias}\{\widehat{m}(x)\} = E\{\widehat{m}(x) - m(x)\}$. Asymptotically, the bias of the Nadaraya-Watson estimator is

$$\text{Bias}\{\widehat{m}(x)\} = \frac{h^2}{2}\mu_2(K_h)\left\{m''(x) + 2\frac{m'(x)f_x'(x)}{f_x(x)}\right\} + O(n^{-1}h^{-1}) + o(h^2) \quad (3.20)$$

with $\mu_2(K_h) = \int u^2 K_h(u)du$ and the variance is

$$\text{Var}\{\widehat{m}(x)\} = \frac{1}{nh}\frac{\sigma^2(x)}{f_x(x)}\int K_h^2(u)du + O(n^{-1}h^{-1}). \quad (3.21)$$

Equations (3.20) and (3.21) show that a reduction in the bandwidth h leads to decreasing bias, but increasing variance. Therefore, determining the optimal bandwidth represents a trade-off between estimation bias and variance.[33,34]

Bandwidth selection is an important step in nonparametric estimation. Bandwidth selection rules are typically based on distance measures such as the mean integrated squared error. The calculation of distance measures is often not straightforward, since they are based on unknown quantities. In practice, *cross validation* and *penalising techniques* have been developed to replace or to estimate the unknown quantities.[35]

Fengler et al. (2003) suggest to determine the optimal bandwidths by minimising the penalisation function

$$pf(h_1, h_2) = n^{-1}\sum_{i=1}^{n}\{\sigma_i - \widehat{\sigma}_{h1,h2}(M_i, T_i)^2 \times \Xi(n^{-1}h_1^{-1}h_2^{-1}K_{h1}(0)K_{h2}(0))\} \quad (3.22)$$

where $\Xi(u) = \exp(2u)$ denotes the Akaike function.[36] Based on this approach, they calculate the optimal bandwidths for each day in their sample. According to Härdle (1990) alternative penalising functions asymptotically deliver the same

[33]Furthermore, the Nadaraya-Watson estimator is consistent, provided that certain regularity conditions are fulfilled. See Fengler (2004), p. 101.
[34]See Schnellen (2007), pp. 11-12.
[35]See Fengler (2004), pp. 104-106.
[36]See Fengler et al. (2003), p. 198.

optimal smoothing parameters, as they share the same first order expansion. In addition, Härdle (1990) demonstrates that the above approach is well suited for bandwidths selection.[37]

The next Section introduces the family of local polynomial kernel estimators, which includes the Nadaraya-Watson estimator as a special case.

Local Polynomial Smoothing

While the Nadaraya-Watson estimator locally fits a constant to the data, local polynomial kernel estimators extend this approach to fitting polynomials in the local neighbourhood. The principle of local polynomial smoothing is to locally approximate the unknown function m using a Taylor series of order p.[38] The local polynomial can be estimated by solving the quadratic optimisation problem

$$\min_{\beta \in \mathbb{R}^{p+1}} \sum_{i=1}^{n} \{y_i - \beta_0 - \beta_1(x - x_i) - \ldots - \beta_p(x - x_i)^p\}^2 \mathbf{K_h}(x - x_i) \tag{3.23}$$

where $\beta = (\beta_0, \ldots, \beta_p)^\top$. The solution of this optimisation problem is given by the usual weighted least squares estimator, where the weighting matrix is equal to $\mathbf{K_h}(x - x_i)$.[39]

In practice, local polynomials of lower order (typically p is 0,1,2, or 3) have been used.[40] Ruppert and Wand (1994) demonstrate that the order of local polynomial estimators should be odd, as they outperform even order polynomials.[41] In particular, the local linear estimator obtained for order $p = 1$ is frequently applied, as its asymptotic bias disappears for a linear function m.[42] Local polynomial smoothing

[37] See Härdle (1990), pp. 165-173.
[38] It is assumed that the function m has continuous derivatives up to order p.
[39] See Schnellen (2007), p. 10.
[40] See Hurvich et al. (1998), p. 272.
[41] See Ruppert and Wand (1994).
[42] The asymptotic bias of the local linear estimator is given by $\text{Bias}\{\widehat{m}(x)\} = \frac{h^2}{2}\mu_2(K)m''(x) + o(h^2)$. See Fengler (2004), p. 103.

is, for instance, used by Aït-Sahalia and Duarte (2003) and Corradi et al. (2012).[43,44] Having described the parametric and nonparametric methods suggested for smoothing implied volatilities, the next Section compares the two approaches.

3.1.4. Comparison of Parametric and Nonparametric Smoothing Methods

The results of the empirical studies regarding the stylised facts of implied volatilities presented in Chapter 2 show that the IVS is not flat across moneyness and maturity. Moreover, the findings provide strong evidence that the IVS changes over time. Therefore, functional flexibility is necessary to account for these features. In particular, this is important when the sample period contains turbulent market periods (e.g., the financial crisis 2008), which may change economic structures.

The recommendations provided in the literature regarding which smoothing method delivers the most appropriate depiction of the IVS are mixed. Walter (2008) argues that most smoothing methods deliver similar results with respect to capturing the characteristic features of the IVS. However, as polynomial smoothing is stable, easy to implement and allows for parsimonious modelling (e.g., by quadratic polynomial smoothing), he prefers this parametric approach.[45] In contrast, Fengler (2004) applies different nonparametric estimators, such as the Nadaraya-Watson estimator and local polynomial smoothing, to DAX options data and suggests that local polynomial smoothing is superior. His perspective of view is that the Nadaraya-Watson estimator can lead to biases at the ends of the implied volatility function because of the unequal distribution of implied volatilities. According to Fengler (2004), this effect does not occur in the same manner if local polynomial smoothing is used.[46]

[43] In addition, semiparametric methods, which combine parametric and nonparametric terms, have also been proposed. Based on a general specification test, Aït-Sahalia et al. (2001) find that a semiparametric model provides a statistically satisfactory description of S&P 500 implied volatilities. See Aït-Sahalia et al. (2001), pp. 385-389.

[44] Schönbucher (1999) and Ledoit et al. (2002) derive such conditions for a single implied volatility and an implied volatility surface.

[45] See Walter (2008), p. 32.

[46] See Fengler (2004), p. 123.

As no method is clearly preferred in the literature, the choice of smoothing method in this study is based on the following argumentation.

The nonparametric approach is employed in this study to smooth DAX implied volatilities, as it provides more flexibility than parametric methods. In addition, there is no financial model capable of fully explaining the observed pricing biases that can be used as reference for a parametric model. As DAX options and futures are liquid instruments, a large sample size is available, which is a prerequisite for the use of nonparametric methods.[47] Of the class of nonparametric estimators, the Nadaraya-Watson estimator is selected in this study.[48] While local polynomial estimators should generally be preferred to the Nadaraya-Watson estimator,[49] I argue, in line with Aït-Sahalia et al. (2001), that the existing sample size in this study is sufficient to deliver similar results.[50] Furthermore, the Nadaraya-Watson estimator has the advantage that it can be easily extended to multidimensional smoothing tasks, which is necessary to smooth the IVS across moneyness and maturity.[51]

After the technical introduction of the Nadaraya-Watson estimator, the next Section contains a description of the underlying data and their preparation. Moreover, the calculation of arbitrage-free DAX implied volatilities based on the BS model and the application of the Nadaraya-Watson estimator to construct the DAX IVS are presented.

[47] See Garcia et al. (2010), p. 531.
[48] Following Fengler et al. (2003) the Nadaraya-Watson estimator with a quartic kernel is used. Note, that the smoothing results are usually more influenced by the selection of the bandwidth parameters than the choice of the kernel. See Behr (2005), p. 61.
[49] See the previous Section.
[50] See Aït-Sahalia et al. (2001), p. 75.
[51] Note, that Horn (2012) also applies the Nadaraya-Watson estimator to smooth DAX implied volatilities.

3.2. Introduction to the Data

3.2.1. Market Structure and Products of the EUREX

The European Exchange (EUREX) was established in September 1998 by the merger of the DTB and the Swiss Options and Financial Futures Exchange (SOFFEX). In 2013, with 2,191 billion traded contracts, the EUREX was the third largest derivatives exchange in the world.[52] The EUREX provides a broad range of products, such as derivatives on interest rates, equities, equity indices, exchange-traded-funds, volatility indices, and credit indices. The most frequently traded options on the EUREX in the year 2012 were options on the EURO STOXX 50 index, at nearly 280.6 million contracts. The second most frequently traded options in 2012 were Deutscher Aktienindex (DAX) options at approximately 51.6 million contracts.[53]

The underlying asset for DAX options, the DAX, was introduced on June 23rd, 1988 and contains the 30 largest and most actively traded German stocks. The base date for the index is December 30th, 1987 where the starting point of the index was set at 1000 points. The stocks are listed in the *Prime Standard* of the Frankfurter Wertpapierbörse. The DAX is constructed as a price index and a capital-weighted performance index where dividends are reinvested. The underlying stock price index for DAX options and DAX futures is the DAX performance index. The index is updated every second using prices from the electronic trading system of the Deutsche Börse AG for the spot market that is called Exchange Electronic Trading (Xetra).[54]

[52]Only the Chicago Mercantile Exchange Group and the Intercontinental Exchange Group reported a higher contract volume. The Futures Industry Association provides the ranking. It is based on the number of futures and options traded and/or cleared in the markets. See Acworth (2014), p. 22.
[53]See EUREX (2013), pp. 65-67.
[54]See Hafner (2004), pp. 73-74

Trading on the EUREX takes place on a completely electronic trading and clearing platform.[55] The trading phase starts at 8:50 a.m. and ends at 5:30 p.m.[56] Options on the DAX are European-style options that can only be exercised on the third Friday of the contract month or on the day before in case of holidays. DAX options are always settled in cash and payable on the first exchange day after the final settlement day. The minimum tick size is 0.1 index points. As DAX options have a contract value of 5 EUR per index point, this corresponds to 0.50 EUR. DAX options are available for the three nearest calendar months, the next three months of the quarterly cycle March-June-September-December, the following four months of the semiannual cycle June-December, and the next two December months of the following two years.[57] The minimum number of exercise prices depends on maturity. For DAX options with maturities up to 24 months, a minimum set of seven different strike prices is provided whereby three ITM strikes, one ATM strike, and three OTM strikes are available. At least five strike prices can be traded for DAX options that expire in more than 24 months. The minimum sets of strike prices are prescribed for the point of issuance. If the DAX moves such that the minimum set of strike prices is no longer available, new strike prices for existing options series are introduced by the EUREX. In addition, the strike price interval varies with maturity. The strike price interval for options with maturities of up to 12 months is 50 index points, 100 index points for maturities from 13 to 24 months, and 200 index points for maturities greater than 24 months.

In addition to DAX options data, this study also uses DAX futures data. After EURO STOXX 50 index futures, DAX futures are the second most heavily traded futures on the EUREX. The contract value of a DAX future is 25 EUR per index point, and the minimum price change is 0.5 index points, which reflects a value of 12.50 EUR. The contract is available for the three nearest quarterly months of the cycle March-June-September-December. As with DAX options, DAX futures

[55] The following details regarding the contract specifications and trading conditions are taken from EUREX (2011) and EUREX (2012).

[56] In addition, pre-trading begins at 7:30 a.m., and the restricted trading phase ends at 8:30 p.m.

[57] In April 2006 the EUREX introduced DAX options with weekly expiration dates.

expire on the third Friday of the contract month and the (final) cash settlement is determined on the last trading day.

3.2.2. Description and Preparation of the Data

The data set used in this empirical analysis contains all recorded transactions of DAX options and DAX futures traded on the EUREX from January 2002 to December 2009.[58] For each traded contract, the sample comprises its type, price, trading volume, time of settlement, maturity, and strike price. The total sample contains 2034 trading days and 6,904,933 transactions.

Figure 3.1 presents the number of options transactions across different maturity classes. It shows that short-term DAX options are the most liquid options. Further, option liquidity tends to decrease for options with longer maturities. Approximately 56% of the transactions account for options with maturities of less than 30 calendar days, and more than 93% have maturities of less than 180 calendar days.

The distribution of options transactions for different moneyness classes is plotted in Figure 3.2, where moneyness is defined as $\ln(K/S)$. It illustrates that option trading is primarily concentrated on ATM options. Moreover, OTM put (ITM call) options are more frequently traded than ITM put (OTM call) options. In addition to the contract data, calculating implied volatilities based on the BS model requires the price of the underlying asset, here the DAX, and the risk-free interest rate. Thus, in the following, the derivation of both input factors is described.

The interest rate data used to compute the implied volatilities cover daily series of the overnight rate Euro OverNight Index Average (EONIA), the Euro InterBank Offered Rate (EURIBOR) for 1 week and 1, 3, 6, 9, 12 months, and the German government bond rates for 2, 3, 4, 5, and 6 years.[59] To ensure that the maturities

[58] With one exception, the options and futures data for the sample period from 2002 and 2009 were provided by the Karlsruher Kapitalmarktdatenbank (KKMDB). The exception is options data for the year 2008, which are obtained directly from the EUREX.

[59] The interest rate data are taken from http://www.bundesbank.de/Navigation/DE/Statistiken/Zeitreihen.

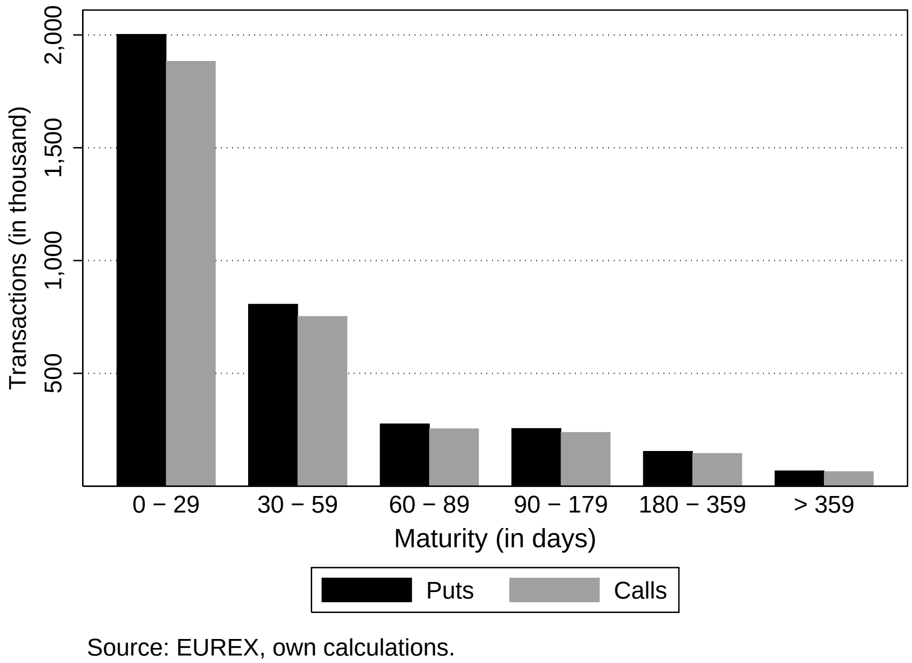


Figure 3.1.: Transactions of DAX options across maturity from 2002 to 2009

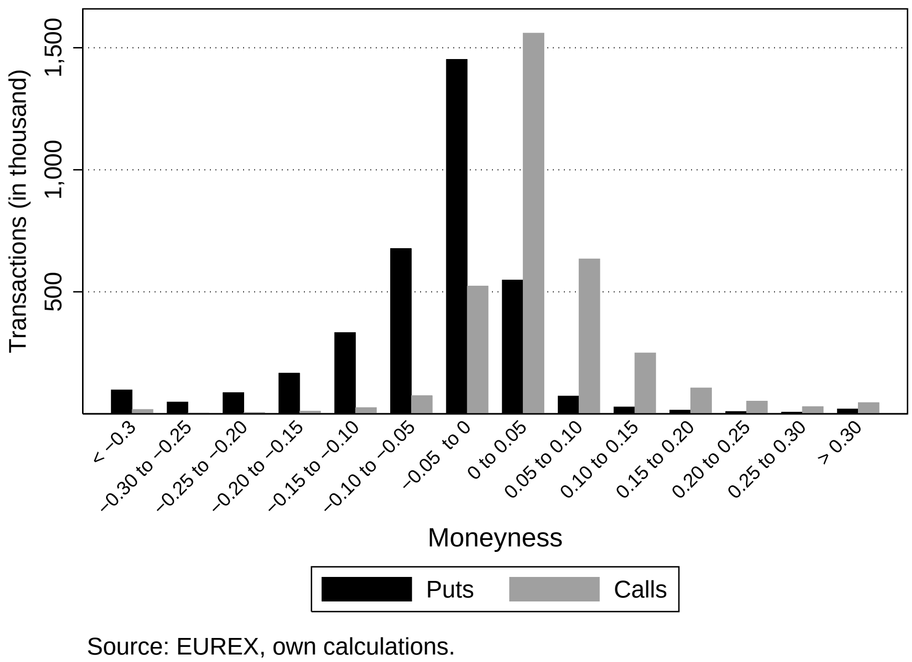


Figure 3.2.: Transactions of DAX options across moneyness from 2002 to 2009

of the interest rate and the option are matched, each interest rate is linearly interpolated from the existing rates enclosing an option's maturity. In the subsequent analysis, continuously compounded interest rates are derived from the nominal rates. Maturity is measured in calendar days and expressed as a fraction of 360 days per year.

The price of the underlying index, DAX, is calculated based on the price of the most liquid DAX futures contract on each day.[60] During the sample period, the most actively traded DAX future on all days was the contract with the shortest maturity. The DAX index level is derived from the current DAX futures price using the cost-of-carry approach

$$S_t = F_t e^{rT} \tag{3.24}$$

where from now on S_t represents the DAX index level at time t, F_t the DAX futures price, r the risk-free interest rate, and T the maturity.[61] To ensure that simultaneous data enter the BS formula, DAX options and futures contracts are matched on each day and hour on a minute-to-minute basis. Therefore, the DAX index level for an option traded on certain day, hour, and minute is computed as the average of all implied S_t obtained from DAX futures that are traded in the corresponding minute. Option transactions are deleted from the data set if no DAX future was traded in the same minute. Implementing a correction scheme for taxes and dividends such as that in Hafner (2004) is not necessary, as income taxation changed in the year 2000.[62] Thus, given the interpolated interest rates and the implied DAX index levels, all information necessary to calculate the implied volatility for each option contract based on the BS model exists.[63]

[60]Liquidity is measured by the number of traded contracts per day.
[61]See Hull (2006), p. 118.
[62]The corresponding Tax Reform Act is the so-called Steuersenkungsgesetz, BGBI. Teil I, Nr. 46. See Fengler (2004), p. 189.
[63]As noted above, the bisection method is used to compute the BS implied volatilities.

3.2.3. Calculation of Arbitrage-Free Implied Volatilities

A necessary step prior to estimating the IVS is to determine whether the option prices satisfy certain no-arbitrage conditions. Cassese and Guidolin (2006) demonstrate that option prices that do not meet the arbitrage conditions can influence the estimation of the IVS and thereby the option pricing mechanism. Thus, many studies employ a filter to eliminate option prices that are incompatible with the absence of arbitrage opportunities. Therefore, similar to Hafner (2004), Fengler (2004), and Gonçalves and Guidolin (2006), among others, this study excludes option prices from the data set that do not fulfil certain arbitrage conditions.

In the literature, the following two no-arbitrage conditions are typically employed: the *upper* and the *lower bound* of the option's price. They are defined as follows:[64]

1. Upper bounds: The purchaser of an American or European call option has the right to buy one share of a stock at strike price K. Thus, the option price c_t can never exceed the current stock price S_t. As the holder of an American or European put option has the right to sell one share of a stock for price K, the price of the option p_t cannot be higher than Ke^{-rT}. The formal conditions are:
   $c_t \leq S_t$ and $p_t \leq Ke^{-rT}$
2. Lower bounds: The value of a European call or put option on a non-dividend-paying asset is never less than its intrinsic value:[65]
   $c_t \geq \max[S_t - Ke^{-rT}, 0]$ and $p_t \geq \max[Ke^{-rT} - S_t, 0]$

The detection of observations that violate one of the above no-arbitrage conditions can be affected by the amount of transaction costs. However, transaction costs are difficult to estimate due to different components, variation across time, and

[64] In addition to these standard no-arbitrage conditions, Cassese and Guidolin (2006) test for a list of further rational pricing bounds, e.g., put/call parity, box spreads, maturity spreads, etc. (see Cassese and Guidolin (2006), p. 155). These more advanced boundaries are not examined, as they require further information on transaction costs.

[65] See Hull (2006), pp. 209-211.

transaction size. Furthermore, they differ among option traders.[66] Therefore, they are not considered.[67]

In total, 34,709 observations are eliminated from the data set due to violations of the above arbitrage bounds. Most of these observations that fail to satisfy the boundary conditions are associated with short-maturity options (with up to 30 days to maturity). Moreover, if options that violate the lower bound condition are considered, the deviation between the option price and the intrinsic value is often low and less than the typical bid-ask spread. This supposes that many violations are generated by transaction costs. In addition, a further 489 observations are excluded because the implied volatilities computed by the bisection-method are not reasonable. After the exclusions, the data set contains 6.87 million option contracts, respectively 99.5% of the total observations.

As the sample period covers volatile market periods, it is reasonable to examine whether the number of violations of the arbitrage bounds is related to the market phase. This question is investigated by Evnine and Rudd (1985), who consider American options on the S&P 100 index traded on the CBOE and options on the Major Market Index traded on the American Stock Exchange between June 26th and August 30th, 1984. They find that all call quotations that do not satisfy the intrinsic-value bound occurred under volatile market conditions in early August 1984. However, while on average, the DAX index options used in this study recorded more violations for the turbulent years 2002, 2003, and 2008, the arbitrage bounds were also violated in more quiet market periods. Furthermore, the violations of the boundaries in the year 2008 were divided nearly evenly between the first and the second half of the year. Thus, the collapse of Lehman Brothers, which occurred in the second half of 2008 and represents a major event of the financial crisis, does not seem to considerably influence the number of arbitrage violations. However, overall this indicates that the number of violations is driven by market phases.

[66] See Mittnik and Rieken (2000), p. 264.
[67] This procedure agrees with Brunner and Hafner (2003), Hafner (2004), and others.

In addition to the validation of the arbitrage bounds, several data filters are used to derive the final data sample. In a first step, DAX options with maturities of less than 5 days are deleted from the data set, as the time premium of these options is very low and this complicates the calculation of implied volatilities.[68] Second, due to liquidity concerns, DAX options with a remaining time to maturity of more than 450 days are not considered.[69] Third, the high number of intraday transactions, misprints, non-synchronous data, data recording problems, and other market imperfections could lead to extreme outliers.[70] Therefore, referring to the related literature, option contracts with implied volatilities below 5% and larger than 120% are deleted.[71] Having applied these filters, the sample is reduced to approximately 5.97 million contracts. Based on the above described two-dimensional Nadaraya-Watson estimator, the IVS is constructed for each day in the sample.[72]

3.2.4. Volatility Regimes

Before analysing the stylised facts of the DAX IVS, the volatility regimes that occurred during the sample period are discussed. Figure 3.3 depicts the time series plots of DAX daily closing prices and DAX implied volatilities from 2002 to 2009. It reveals two extremely volatile market periods at the start (mid-2002 to mid-2003) and at the end of the sample period (the 2008 financial crisis). During both periods, the DAX declined dramatically and DAX implied volatility increased considerably.

[68] In this case, the computation of implied volatilities is highly sensitive to measurement errors. See Hafner (2004), p. 90.

[69] This restriction also follows from the following analysis of the IVS, which concentrates on DAX implied volatilities for maturities of up to 12 months.

[70] For instance, a transaction that is erroneously entered by an option trader, a mistrade, can be annulled by the EUREX under certain conditions.

[71] In comparison to the related literature the implemented filter is mild. For instance, Fengler (2004) deletes observations with implied volatilities below 4% and above 80% and Herrmann (1999) only uses implied volatilities below 40%.

[72] In order to reduce computational burden, the bandwidth in the moneyness dimension, h_1, is set to 0.05 and the bandwidth in the maturity dimension, h_2, to 0.25 for each day. Both values are determined based on visual examination of the resulting observations and are close to the average optimal bandwidths ($\overline{h}_1 = 0.03$ and $\overline{h}_2 = 0.23$).

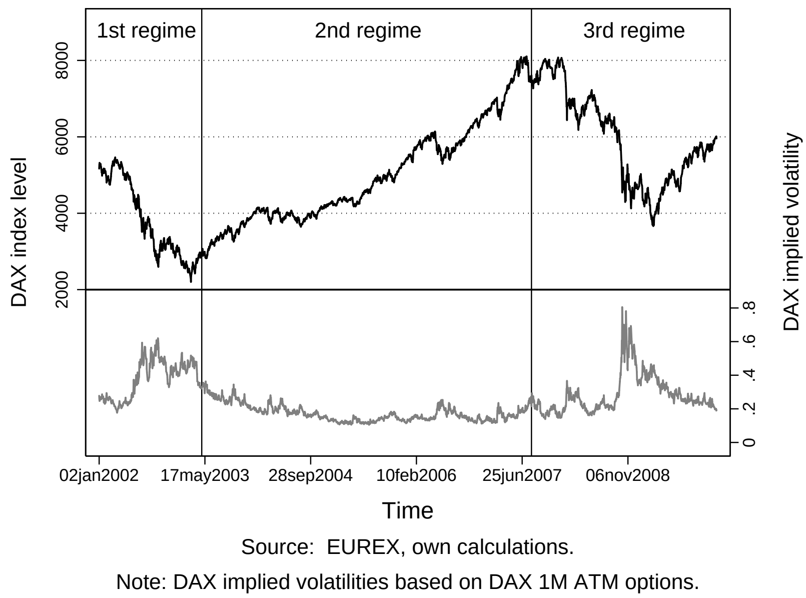


Figure 3.3.: DAX index level and DAX implied volatilities from 2002 to 2009

In the following, certain major events that influenced the largest changes in the DAX during the sample period are presented.

The extreme decline in the DAX in 2002 was attributed to increasing investor concerns regarding an upcoming recession in the US economy. The further decline in the DAX index in 2003 was caused by resurfacing doubts concerning the duration of the Iraq war. Following this turbulent market phase, in March 2003 the DAX began a long, relatively stable upturn that ended in August 2007 with the beginning of the financial crisis. In the literature, August 9th, 2007 is often cited as the starting point of the financial crisis. On this day, the French bank BNP Paribas temporarily halted redemptions from three investment funds that were invested in subprime mortgage debt, as a reliable valuation of the funds was no longer possible. Thereafter, financial institutions worldwide questioned the values of a variety of collaterals. Faced with growing market uncertainty, financial institutions hoarded cash, inter-bank lending dried up, and many market players were faced with severe liquidity constraints. Therefore, the supply of short-term funds diminished and the overnight interest rates in Europe shot up. To provide liquidity, the European Central Bank decided to arrange the largest short-term liquidity injection in its nine-year history. Distress in the credit market also influenced the stock and derivatives markets, which can be observed in the increased DAX implied volatilities in August 2007 (see Figure 3.3).

These initial disturbances persisted until the middle to end of September 2007. The next volatility peak was on January 21st, 2008 when the DAX-Volatilitätsindex (VDAX) increased to nearly 30% and the DAX lost temporarily more than 7% of its value. This was the largest loss since September 11th, 2001.[73] The financial crisis peaked between September and November 2008. On September 15th, 2008 Lehman Brothers filed for bankruptcy protection, and in the following days the VDAX approached the 30% level. On October 16th, 2008, the VDAX reached its highest level in the past ten years at 74%. The DAX lost more than 40% of its value between September 15th and November 21st, 2008. At the end of year 2008, the VDAX level returned to lower levels and fell below 40%.

In the following, different subsamples are defined to analyse the behaviour of the DAX IVS during different volatility regimes (see also Figure 3.3 for the definitions of the volatility regimes). Given the above mentioned events and course of DAX implied volatility over the full sample period, the sample is divided into three subsamples. The first sample (or first regime) considers the turbulent market phase at the beginning of the sample period from January 2nd, 2002, to May 2nd, 2003. As no clear external event marks the start of the long, stable market upturn from the spring of 2003 to mid-2007, May 2nd, 2003, when the DAX returned to an index value of 3000 is selected. The end of the second subsample (or second regime) is August 8th, 2007, which is the day before the above-cited starting point of the financial crisis. Thus, the second subsample comprises the time period from May 5th, 2003, to August 8th, 2007 and the last subsample (third regime) the period from August 9th, 2007 to December 30th, 2009.

[73] The explanation of *Black Monday* is unclear. First, market participants feared an upcoming US recession, which induced panic in the markets (see Landler and Timmons (2008)). Further, it become known after a few days that the French bank Société Generale closed out high positions between January 21st and January 23rd, 2008, which were created by a trader employed at the company. According to the bank, these positions were fraudulent and led to a loss of approximately 4.9 billion EUR. See Viscusi and Chassany (2008).

3.3. Stylised Empirical Facts of the DAX IVS

As major stock markets crashed in October 1987, the volatility smile is typically more pronounced and downward sloping. In particular, short-term OTM options exhibited higher implied volatilities than ATM options.[74] Therefore, an extreme market downturn can permanently change the option pricing behaviour of market participants.[75] Empirical studies that consider the effect of a stock market crash on implied volatilities include Schwert (1990), Bates (1991), Bates (2000), Constantinides et al. (2009), and Schwert (2011). All of these studies primarily consider the US options market by using options on the S&P 500 index and/or S&P 500 futures. In contrast, relatively little is known about the German options market.

While there are several studies in the literature on the characteristics of DAX implied volatilities during normal market conditions, [76], few empirical studies compare the behaviours of DAX implied volatilities before, during, and after financial crises.[77] Therefore, the following Section analyses whether the crises in 2002/2003 and 2008 influenced DAX implied volatilities. The results are relevant for the prediction of DAX volatility. Because DAX volatility forecasts in this study are constructed based on DAX implied volatilities for a sample period that includes two dramatic stock market declines, it is of particular interest whether the pattern of implied volatilities changed due to these financial crises. If systematic structural changes occurred, the option pricing model should explicitly account for these changes or provide sufficient flexibility to rapidly adapt to these changes. Unsurprisingly, the inclusion of observations around a market crash affects the estimation of the IVS. For instance, Harvey and Whaley (1992) report that these observations can considerably influence the estimation of a regression model in which implied volatility changes

[74] See Chapter 2.5.1.
[75] See Benzoni et al. (2011), p. 552.
[76] See Chapter 2.5.
[77] An exception is Trautmann and Beinert (1995), who report that the inferred risk neutral skewness from German stock options increased after the 1987 crash.

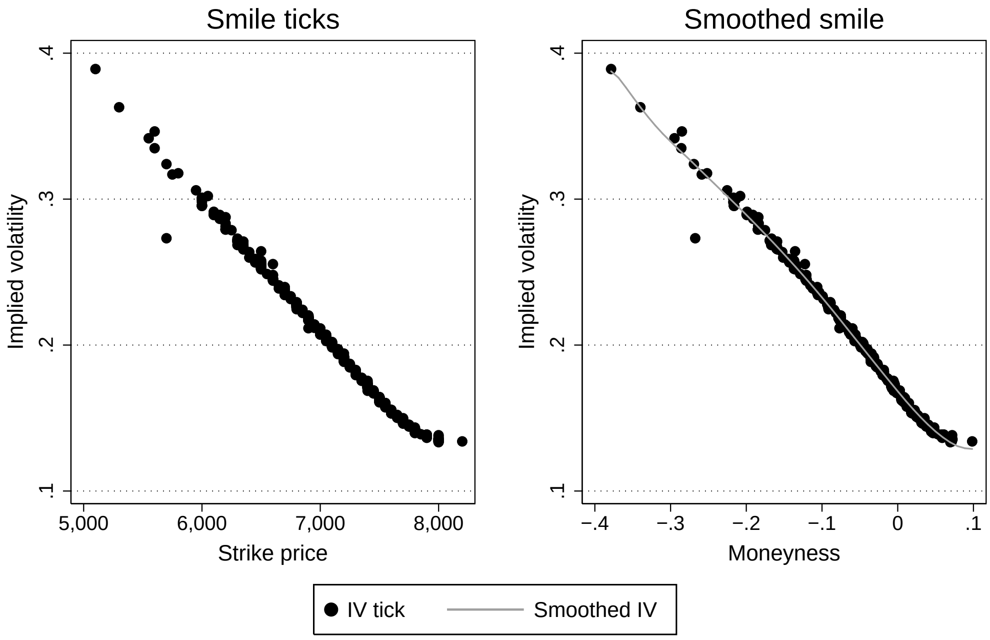
are explained by different factors.[78] In the following, the empirical regularities of the IVS and their variations are described for the above-defined subsamples.

3.3.1. DAX Volatility Smiles

A typical picture of the DAX volatility smile is provided in Figure 3.4. In the left panel of Figure 3.4, implied volatilities are plotted against the strike price dimension for DAX options with 44 days to maturity that were traded on May 2nd, 2007. It can be observed that the implied volatilities are not constant across strike prices, but rather form a downward-sloping skew. This implies that DAX OTM puts and DAX ITM calls trade at higher prices than DAX ATM options with the same maturity. The volatility smile appears as a black line in Figure 3.4, as implied volatilities are not available for all strike prices due to standardised contract specifications and trading activity. To analyse volatility smiles over longer time periods, implied volatilities are typically plotted against moneyness. In so doing, the smile is not influenced by large changes in the underlying asset price. Another advantage of using moneyness is that it acts as a quasi smoothing method.[79] A smoothed volatility smile where the implied volatilities are estimated by the Nadaraya-Watson estimator is depicted in the right panel of Figure 3.4.

Next, it is investigated whether the DAX volatility smile varies across maturity. The left panel of Figure 3.5 presents DAX volatility smiles for different maturities that were observed on May 2nd, 2007. It demonstrates the smile pattern exists for all considered maturities and that the smile tends to flatten out for long-term options. Moreover, considering OTM puts/ITM calls, the implied volatilities of short-term options seem to be higher than those of long-term options with the same moneyness level. The right panel of Figure 3.5 contains DAX volatility smiles for selected maturities on October 16th, 2008, when implied volatility reached its highest

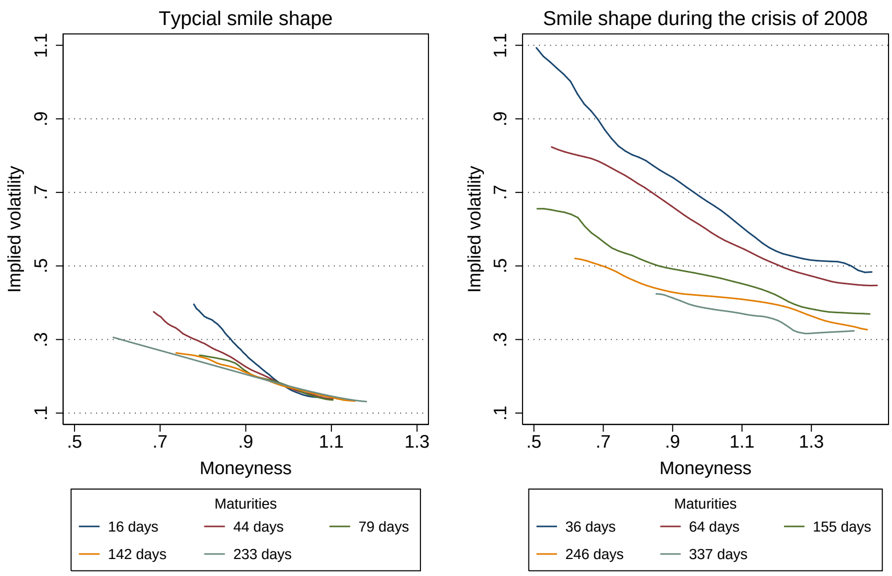
[78]See Harvey and Whaley (1992), pp. 60-61.
[79]See Fengler (2004), p. 24.



Source: EUREX, own calculations.

Note: Implied volatilities (IV) on May, 2nd 2007 based on DAX options with 44 days to maturity.

Figure 3.4.: DAX implied volatility smiles



Source: EUREX, own calculations.

Note: DAX implied volatilities on May, 2nd 2007 (left panel) and on Oct, 16th 2008 (right panel).

Figure 3.5.: DAX volatility smiles for different maturities

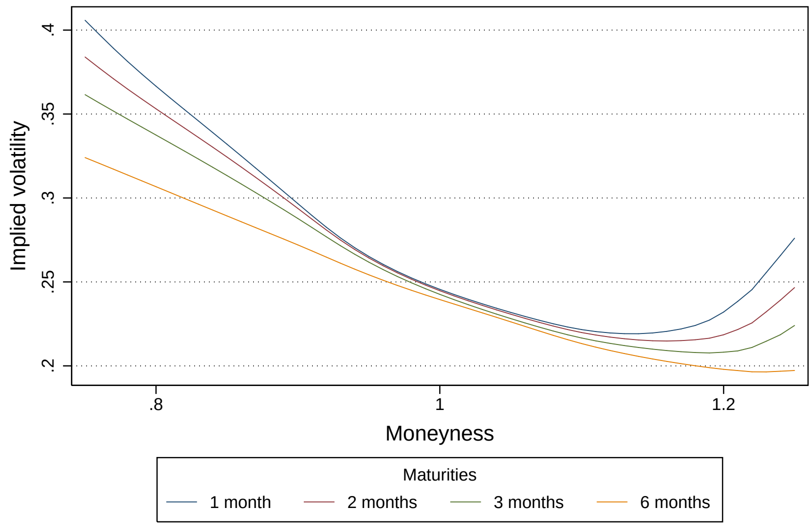
level in ten years. It suggests that while the smile pattern was still present for all maturities, the smiles shifted to higher levels.

Figure 3.6 shows that, on average, a DAX volatility smile, respectively skew, existed throughout the sample period. In particular, it demonstrates that on average, a volatility smile can be observed for short-term options (with maturities of up to 3 months) and a volatility skew for long-term options (with a maturity equal to 6 months). Moreover, the figure indicates that average implied volatility takes its minimum above the ATM level. Further, the average volatility smile of DAX short-term options lies above the average volatility skew of DAX options with 6 months to expiry. Hafner (2004) provides similar results regarding the shape of the DAX volatility smile for a sample period from 1995 to 2002. First, he also finds that the average implied volatilities of DAX OTM puts/ITM calls with 1 month to expiry were higher than those of DAX options with the same moneyness level but 3 or 5 months to expiration. Nevertheless, his findings differ from the results of this study in some respects. Whereas he reports that the volatility smiles of DAX options with 1, 3, and 5 months to maturity intersect slightly below the ATM level, this study finds that the average estimated smile of short-term DAX options resembles a tangent to the smile curves of long-term DAX options. As a consequence, he finds that the implied volatilities of DAX ITM puts/OTM calls with longer maturities (more than 3 or 5 months to maturity) exceeded the implied volatilities of DAX short-term options (1 month to maturity) with the same moneyness level. Furthermore, in contrast to the findings of this work presented above, he reports that a volatility skew can generally be observed for DAX short-term options during the period from 1995 to 2002.[80,81]

The average DAX volatility smile for the above-defined volatility regimes and different maturities is presented in Figure 3.7. While a DAX volatility smile was generally documented for short-term options in the total sample, a skew pattern occurred dur-

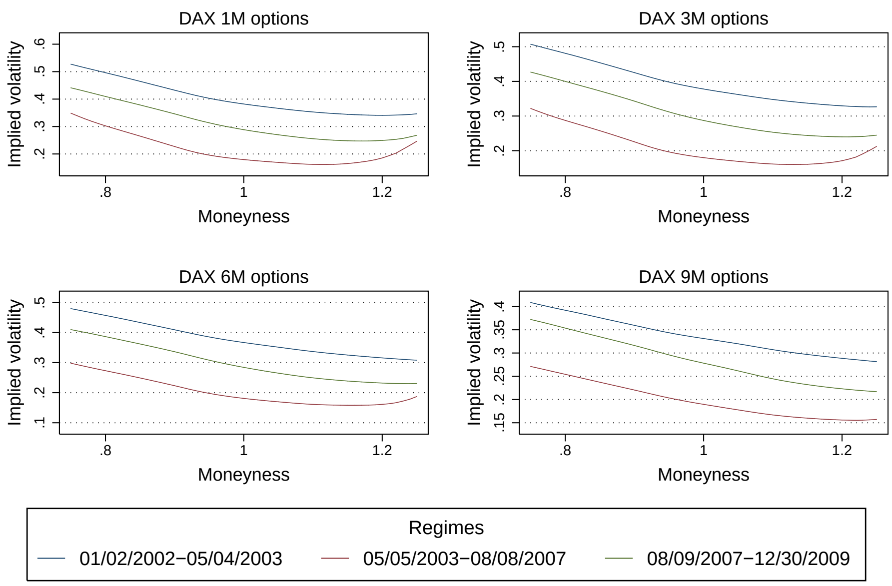
[80] See Hafner (2004), p. 96.

[81] He also observes volatility smiles in the sample but mentions that this is not the typical smile pattern throughout the sample. See ibid., p. 95.



Source: EUREX, own calculations.
Note: Implied volatilities based on DAX options.

Figure 3.6.: Average DAX volatility smiles from 2002 to 2009



Source: EUREX, own calculations.
Note: Implied volatilities based on DAX options.

Figure 3.7.: Average DAX volatility smiles for different volatility regimes

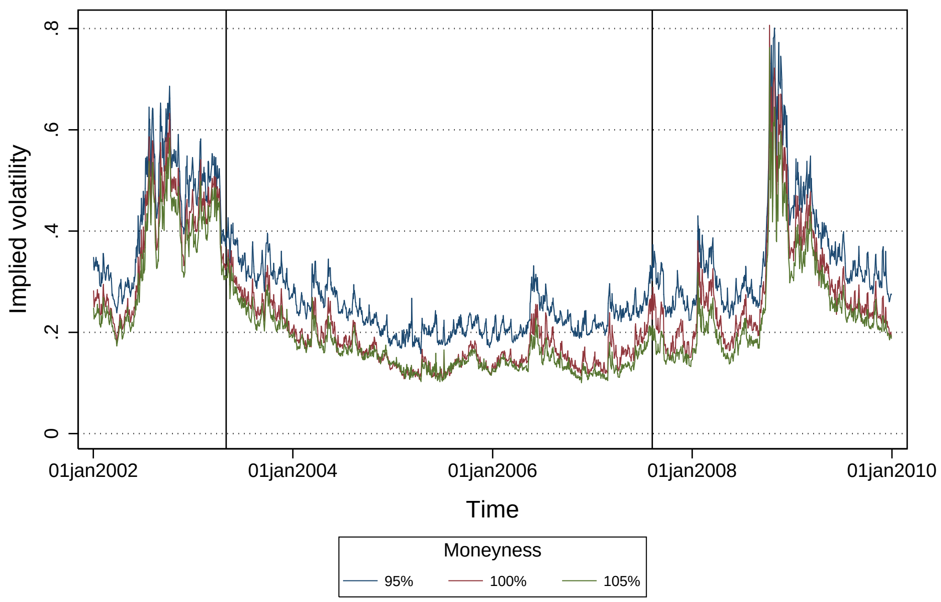
ing the first subsample. Further, the smile pattern of the DAX implied volatility curve during the third subsample is also not particularly pronounced. Therefore, the different findings obtained by Hafner (2004) and this study concerning the shape of the DAX volatility smile for short-term options may be due to the long, stable market upturn included in the period considered in this study. Additionally, Figure 3.7 confirms the above assumption that DAX implied volatilities move upwards (downwards) during turbulent (stable) market periods. This behaviour can be observed for all DAX implied volatilities across all maturities and moneyness levels. However, although the DAX level increased in 2009 and DAX implied volatilities decreased, they did not revert back to their pre-crash levels.[82] Therefore, market participants traded DAX short-term options after the financial crisis at higher prices than before. However, the post-crash sample period is too brief to ultimately conclude whether this is a temporary or permanent effect. This time series variation in short-term DAX implied volatilities for different degrees of moneyness can be observed in Figure 3.8. Further, the figure demonstrates that DAX implied volatilities for different moneyness levels typically moved parallel to one another during the sample period.[83]

In the following, the skewness of the DAX volatility smile is examined. For this purpose, three skewness measures are introduced that are similar to the statistics suggested by Wallmeier (2003).[84] The first two skewness measures are defined as the difference between the implied volatilities of two DAX options (with the same maturity but different moneyness levels) divided by the implied volatility of a DAX

[82] This observation refers to DAX short-term options.

[83] The changes in the implied volatilities are highly correlated. For instance, the correlation between the implied volatilities of DAX 1M ATM options and DAX 1M options with a moneyness level of 95% is 95.2%.

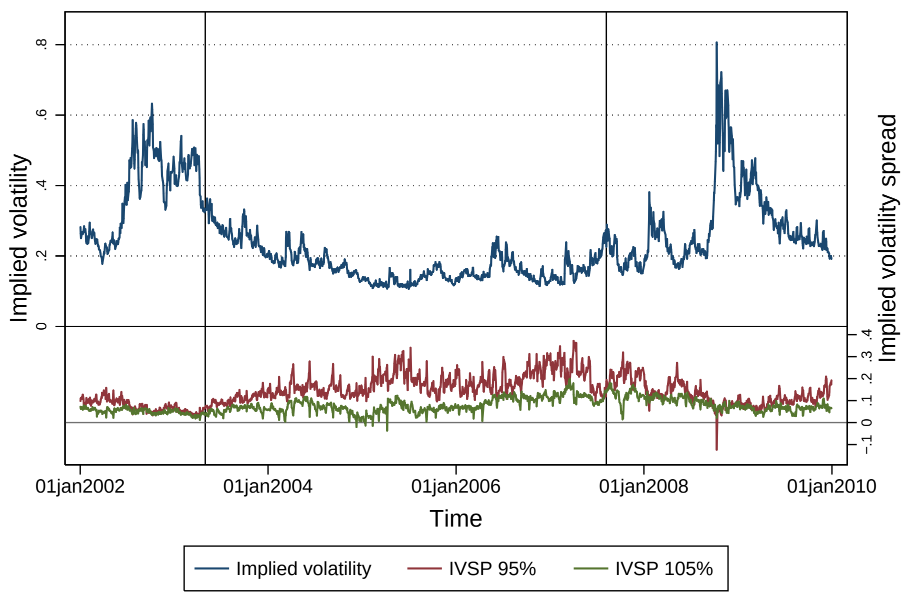
[84] See Wallmeier (2003), p. 190.



Source: EUREX, own calculations.

Note: Implied volatilities based on DAX 1M options.

Figure 3.8.: DAX implied volatilities for different moneyness levels



Source: EUREX, own calculations.

Note: Implied volatilities based on DAX 3M options.

Figure 3.9.: DAX implied volatility and volatility spreads

ATM option. They are called *implied volatility spreads (IVSP)*. Formally, their calculation is based on

$$IVSP_{95\%} = \frac{\widehat{\sigma}_t(M_1,T) - \widehat{\sigma}_t(M_0,T)}{\widehat{\sigma}_t(M_0,T)} \tag{3.25}$$

$$IVSP_{105\%} = \frac{\widehat{\sigma}_t(M_0,T) - \widehat{\sigma}_t(M_2,T)}{\widehat{\sigma}_t(M_0,T)} \tag{3.26}$$

where the moneyness levels are $M_0 = 1$, $M_1 = 0.95$ and $M_2 = 1.05$. An additional skewness measure that covers the sum of both implied volatility spreads is referred to as the total volatility spread (TVSP) and is given by

$$TVSP = \frac{\widehat{\sigma}_t(M_1,T) - \widehat{\sigma}_t(M_2,T)}{\widehat{\sigma}_t(M_0,T)}. \tag{3.27}$$

Positive values of the implied volatility spreads indicate a downward sloping smile curve. Thus, the positive values of the implied volatility spreads in Figure 3.9 reflect that the DAX smile was generally negatively skewed during the sample period.[85] Similar results are reported by Wallmeier (2003) for DAX options traded between 1995 and 2000.[86] He finds that both implied volatility spreads were fully positive and therefore an asymmetric DAX volatility skew existed in the sample period.[87] Moreover, Hafner (2004) finds for the time period from 1995 to 2002 that the slope of the DAX volatility smile was generally negative and steeper for short-term options.[88] These empirical observations agree with the theoretical findings of Rogers and Tehranchi (2010), who show that the implied volatility surface flattens out if maturity goes to infinity.[89]

Interestingly, the average difference between the two implied volatility spreads during the second subsample, which comprises a long, stable market upturn, was higher than in the two other, more volatile subsamples. This can be attributed to a higher

[85] For instance, one exception was observed on October 10th, 2008, where the smile was positively skewed for one day.
[86] Wallmeier (2003) considers DAX options with 45 days to maturity.
[87] See ibid., p. 191.
[88] See Hafner (2004), p. 96.
[89] See Rogers and Tehranchi (2010), pp. 243-247.

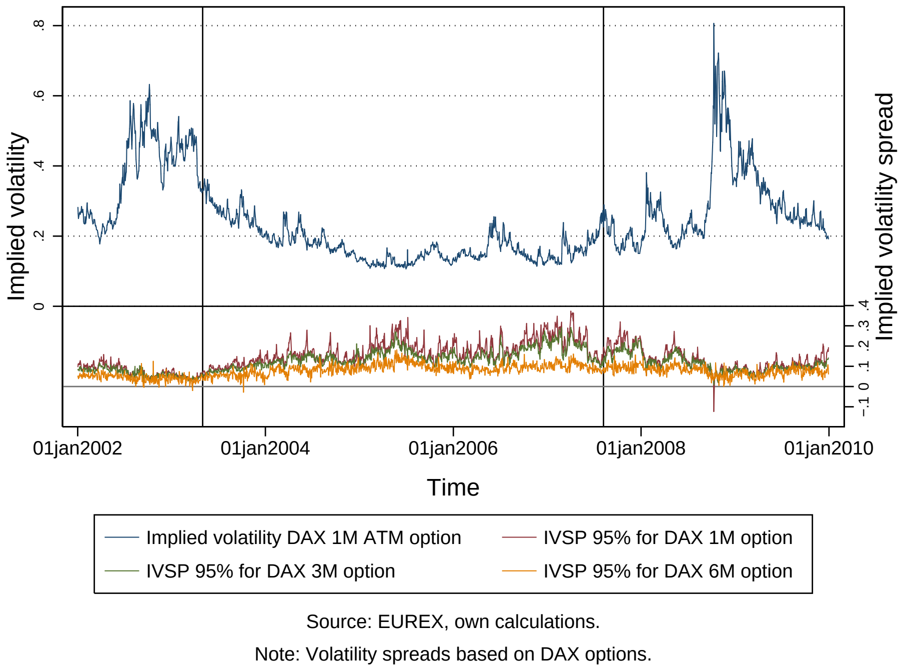


Figure 3.10.: DAX implied volatility and volatility spreads for different maturities

increase in $IVSP_{95\%}$ relative to $IVSP_{105\%}$ in the second subsample.[90] Thus, while the level of DAX implied volatility declined to lower, more normal values, the implied volatility spreads increased. Therefore, although the market was comparatively stable during May 2003 and August 2007, market participants paid higher relative prices for DAX OTM puts providing portfolio insurance. Moreover, according to Figure 3.9, the two implied volatility spreads were more similar in turbulent market periods. When considering the total volatility spread, this study confirms the findings of Wallmeier (2003), who documents that an increase in the implied volatility of DAX ATM options generally leads to a reduction in the total volatility spread.[91]

Figure 3.10 depicts the time series of the implied volatility spreads $IVSP_{95\%}$ for DAX options with 1, 3, and 6 months to maturity. It shows that the implied volatility spread series of DAX options with 1 month, respectively 3 months, seem to be more

[90] A correlation analysis of the changes in DAX ATM implied volatilities and implied volatility spreads reveals the following pattern: whereas a decrease in DAX implied volatility tends to increase the volatility spread $IVSP_{95\%}$, the opposite can be observed for the $IVSP_{105\%}$.

[91] See Wallmeier (2003), p. 193.

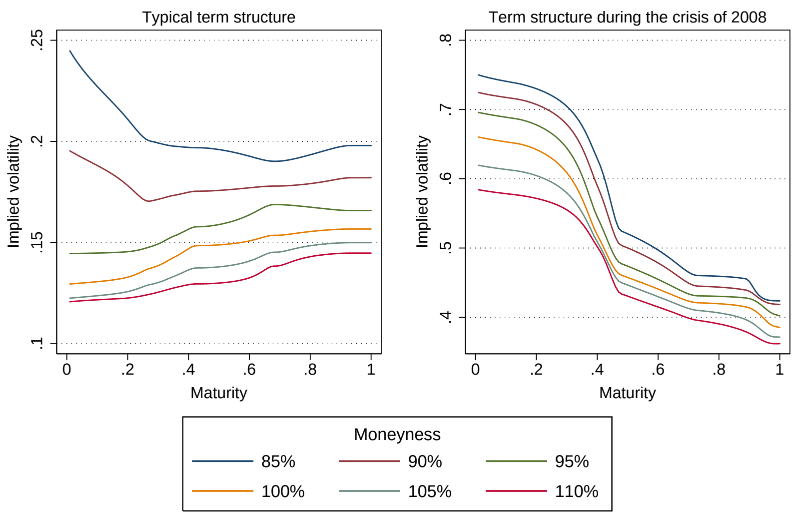
closely related than either series is to the implied volatility spread of DAX options that expire in 6 months. Moreover, the figure illustrates that the slope of the DAX volatility smile changes during the sample period, particularly during volatile market periods. This behaviour can be observed for each of the three considered implied volatility spread series.

Having described the DAX volatility smile, the next Section examines the DAX volatility term structure.

3.3.2. DAX Volatility Term Structures

A typical DAX volatility term structure during the second volatility regime is depicted in the left panel of Figure 3.11. The figure presents the DAX volatility term structure for different moneyness levels that were observed on January 16th, 2006. It indicates that the term structure changes its shape across different moneyness levels. While the volatility term structure for DAX options with moneyness levels of 85% and 90% is U-shaped, it shows an increasing pattern for DAX options with a moneyness level above 95%. In contrast, the right panel of Figure 3.11 illustrates that each DAX volatility term structure for a given moneyness level became downward sloping in an extremely volatile market setting. The figure also demonstrates for both days that, given the same maturity, the DAX implied volatilities seem to increase if the moneyness level decreases.

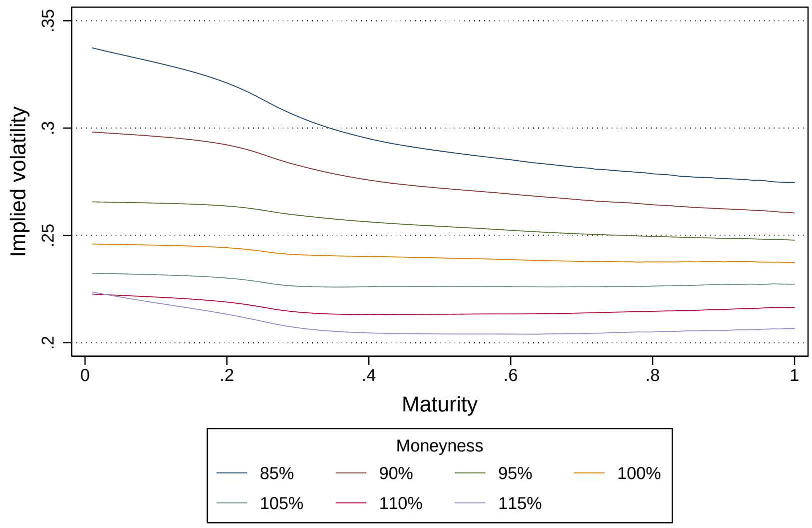
In the following, the average DAX volatility term structure for the total sample across different moneyness levels is examined. Figure 3.12 shows that from 2002 to 2009, a decreasing volatility term structure generally existed for DAX options with a moneyness level below 100%. The term structures for DAX ATM options and DAX OTM call/ITM put options were decreasing for maturities of up to 4 months. For higher maturities, the term structure was nearly flat (in the case of ATM options) or marginally increasing (for options with a moneyness level above 100%). In general, the average term structures of DAX options with lower moneyness levels lie above



Source: EUREX, own calculations.

Note: DAX implied volatilities on Jan, 16th 2006 (left panel) and on Oct, 16th 2008 (right panel).

Figure 3.11.: DAX term structures



Source: EUREX, own calculations.

Note: Implied volatilities based on DAX options.

Figure 3.12.: Average DAX term structures for different moneyness levels

the curves of options with higher moneyness. Hafner (2004) reports that the average term structure of DAX options with a low moneyness level (e.g., 85%) was decreasing between 1995 and 2002. In contrast, he finds an increasing average term structure for DAX ATM options and options with higher moneyness levels. In addition, he observes frequent changes between increasing and decreasing DAX volatility term structures.[92] Wallmeier (2003) confirms the findings of Hafner (2004) by reporting an increasing term structure for DAX ATM options for the period from 1995 to 2000. Fengler (2004) also provides similar results. Moreover, he finds that the average yearly DAX ATM volatility term structures are increasing from 1995 to 2001.[93]

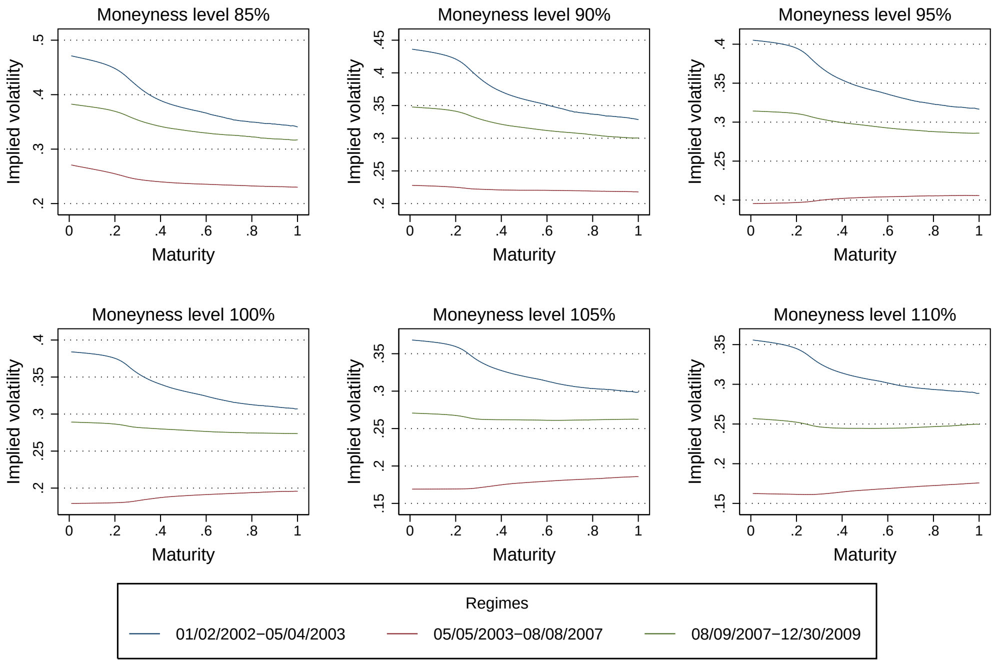
However, Figure 3.13 shows that different DAX volatility term structures can be observed for the volatility regimes. In addition the overall upward shift in DAX implied volatilities, the slope and the curvature of the volatility term structure changed during volatile market phases. Whereas Figure 3.12 depicts a nearly flat term structure of DAX ATM options (with maturities above 4 months), Figure 3.13 provides evidence that this is due to taking the average of an increasing term structure during the second subsample and a decreasing shape in the first and third subsamples (see also the term structures for DAX options with higher moneyness levels). Furthermore, the figure indicates that the curvature of DAX options with higher moneyness levels is more affected by extreme market movements than those of options with moneyness levels below 95%. In addition, Figure 3.14, which depicts the times series of implied volatilities for DAX ATM options with 1, 3, and 6 months to maturity, also indicates that the term structure changed considerably during the sample period and the movements of the implied volatilities series were closely related.[94]

Next, the slope of the DAX volatility term structure is analysed. Similar to the above skewness measures for the volatility smile, three slope measures for the term structure are defined. The first two slope measures of the term structure are equal

[92]While the volatility term structure for DAX ATM options exhibited an increasing shape on 1316 days, the inverse shape could be documented for 623 days in the sample. See Hafner (2004), pp. 95-96.

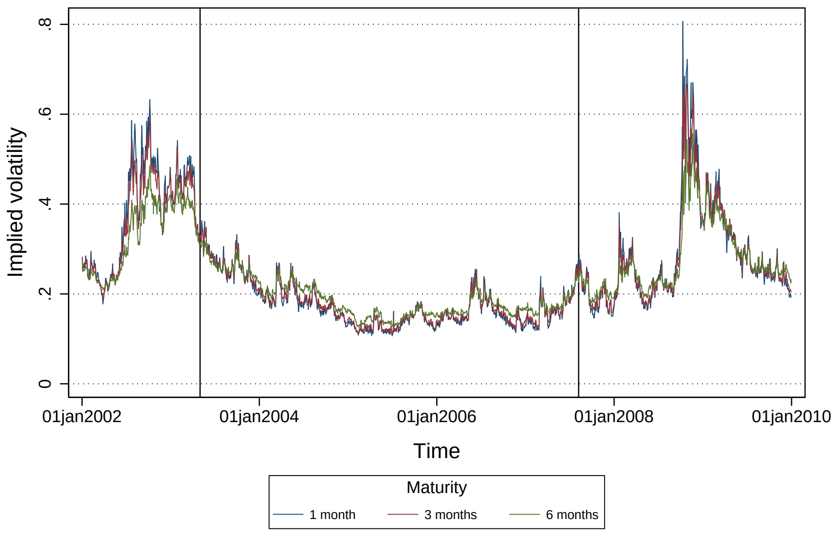
[93]See Fengler (2004), p. 31.

[94]The correlations of the changes in the DAX implied volatilities series are between 0.82 and 0.95.



Source: EUREX, own calculations.
Note: Implied volatilities based on DAX options.

Figure 3.13.: Average DAX term structures for different volatility regimes



Source: EUREX, own calculations.
Note: Implied volatilities based on DAX ATM options.

Figure 3.14.: DAX implied volatilities for different maturities

to the difference between the implied volatilities of two DAX options (with the same moneyness level but different maturities) divided by the implied volatility of a DAX 3M ATM option. Both measures are called *implied volatility term structure spreads (IVTSP)* and are computed by

$$IVTSP_{1-3m} = \frac{\widehat{\sigma}_t(M,T_2) - \widehat{\sigma}_t(M,T_1)}{\widehat{\sigma}_t(M,T_2)} \tag{3.28}$$

$$IVTSP_{3-6m} = \frac{\widehat{\sigma}_t(M,T_3) - \widehat{\sigma}_t(M,T_2)}{\widehat{\sigma}_t(M,T_2)} \tag{3.29}$$

where the maturities are $T_1 = 1\,\text{month}$, $T_2 = 3\,\text{months}$, and $T_3 = 6\,\text{months}$. The third slope measure is defined as the sum of $IVTSP_{1-3m}$ and $IVTSP_{3-6m}$ and is given by

$$IVTSP_{1-6m} = \frac{\widehat{\sigma}_t(M,T_3) - \widehat{\sigma}_t(M,T_1)}{\widehat{\sigma}_t(M,T_2)}. \tag{3.30}$$

Positive (negative) values of the implied volatility term structure spreads reflect an increasing (decreasing) volatility term structure. Figure 3.15 indicates that the slope of the term structure is time-varying. After accounting for the volatility regimes, it seems that the slope of the term structure tends to be positive during market periods with normal volatility levels. In contrast, in turbulent market phases, a negative slope is generally observed. The change in the slope during volatile periods has not been previously documented in the literature for DAX options. For instance, Fengler (2004) reports a more or less flat average term structure for DAX ATM options in 2001 when the dot-com crisis affected stock markets. However, this finding may be the result of averaging the daily term structures over the year 2001.[95] Finally, Figure 3.16 illustrates the time series of $IVTSP_{1-6m}$ for different moneyness levels. It is worth noting that in the second subsample period, the behaviour of the slope measure series differs considerably more across the three moneyness levels than in the other subsamples.

[95] See Fengler (2004), p. 31.

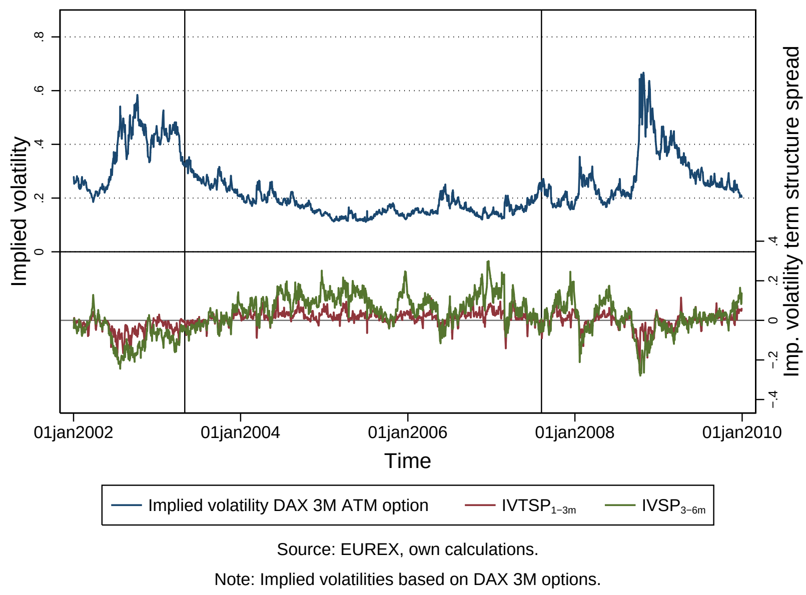


Figure 3.15.: DAX implied volatility and volatility term structure spreads

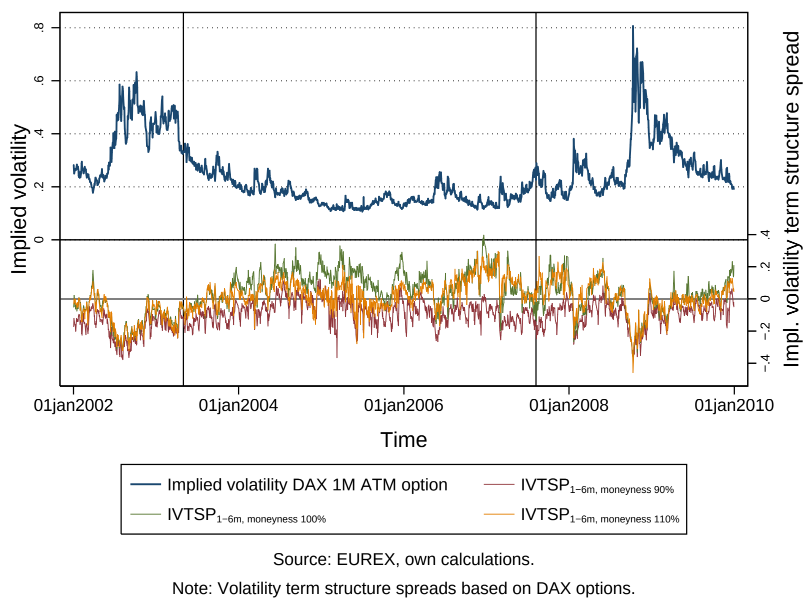


Figure 3.16.: DAX implied volatility and volatility term structure spreads for different moneyness levels

3.3.3. DAX IVS

Having analysed DAX volatility smiles and volatility term structures, this Section presents the DAX IVS, which combines both effects. Despite the in-depth discussion of these effects presented above, a graphical illustration of the IVS is helpful for understanding the simultaneous variation in the implied volatility across both dimensions. As volatility smiles and term structures have been analysed in depth in the previous Sections, the shape of the DAX IVS is only briefly described here.

Figure 3.17 depicts the average DAX IVS for the complete sample period from 2002 to 2009. It shows that the DAX volatility smile is steepest for short-term options and flattens out with increasing maturity. Moreover, as mentioned above, a decreasing DAX term structure can be observed for DAX options with a moneyness level below 100% (see Chapter 3.3.2 for a description of the DAX term structure for options with higher moneyness). As a DAX volatility smile was generally identified for short-term options in the total sample (see Chapter 3.3.1), the average DAX IVS differs from the findings of Hafner (2004) and Fengler (2004).[96] The volatility of DAX implied volatilities is depicted in Figure 3.18. It can be observed that the implied volatilities of DAX short-term options were more volatile than those of options with longer maturities. Fengler (2004) provides similar results.

The shape of the average DAX IVS for the different volatility regimes is presented in Figures 3.19, 3.20, and 3.21. A comparison of the IVS shapes across the regimes emphasises the finding presented above that the DAX volatility smile of short-term options is more pronounced in the second subsample than in the two other subsamples. Finally, Figure 3.22 presents the times series of DAX implied volatilities with different maturities and moneyness levels. It suggests that the series of DAX implied volatilities across maturity and moneyness are highly correlated.

[96]The appendix contains two additional IVS. Figure A.1 illustrates a typical DAX IVS for the second subsample that was observed on May 2nd, 2007. A DAX IVS that occurred on an extremely volatile day, October 16th, 2008, is depicted in Figure A.2.

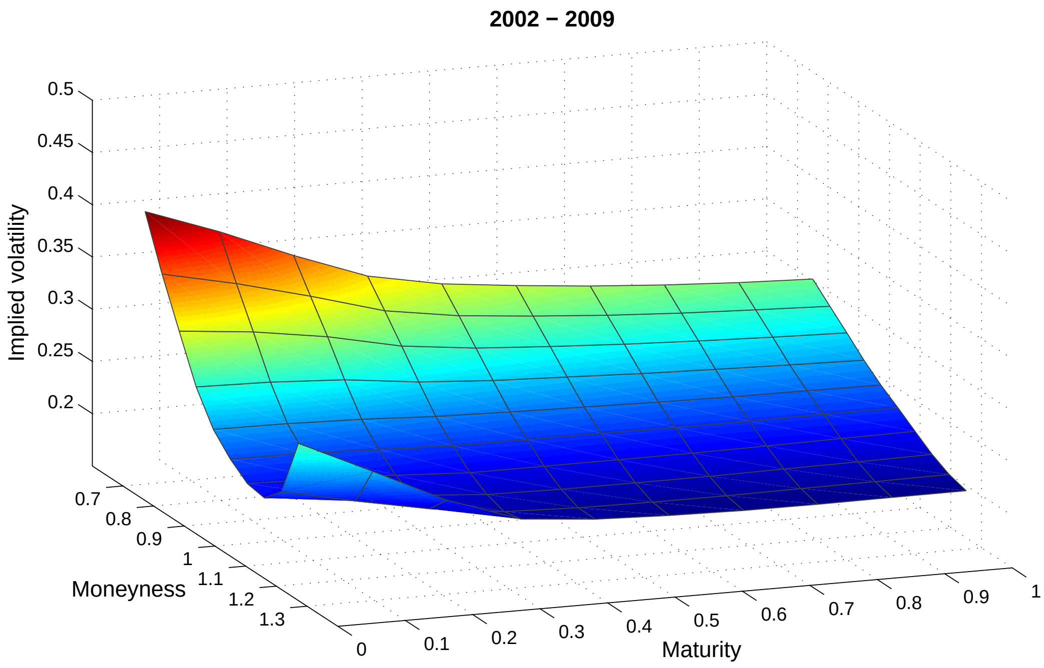


Figure 3.17.: Average DAX IVS for the sample period from 2002 to 2009

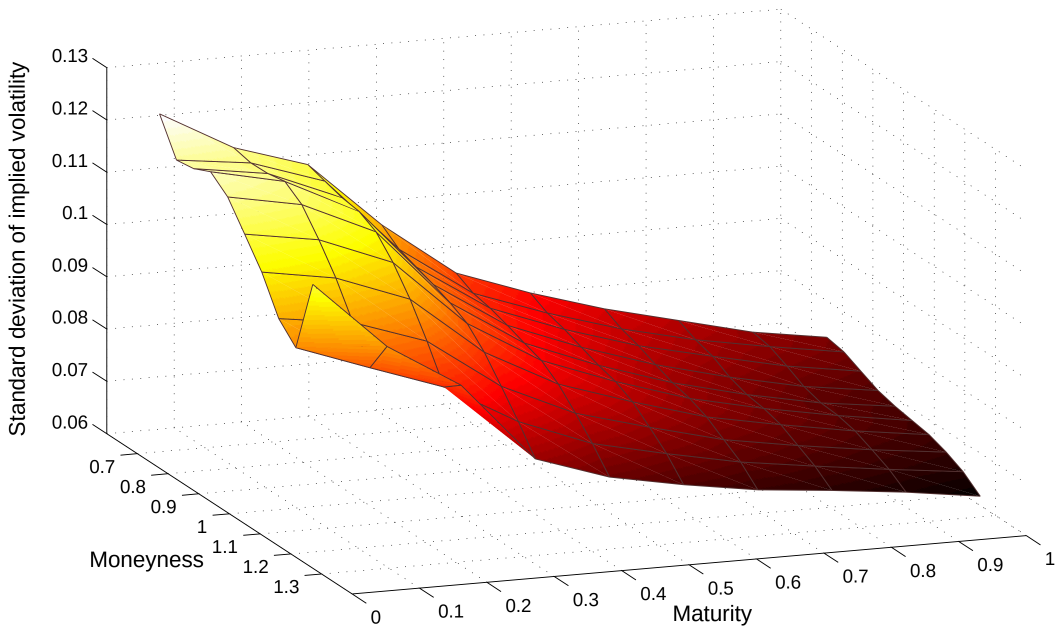


Figure 3.18.: Standard deviation of the DAX IVS for the sample period from 2002 to 2009

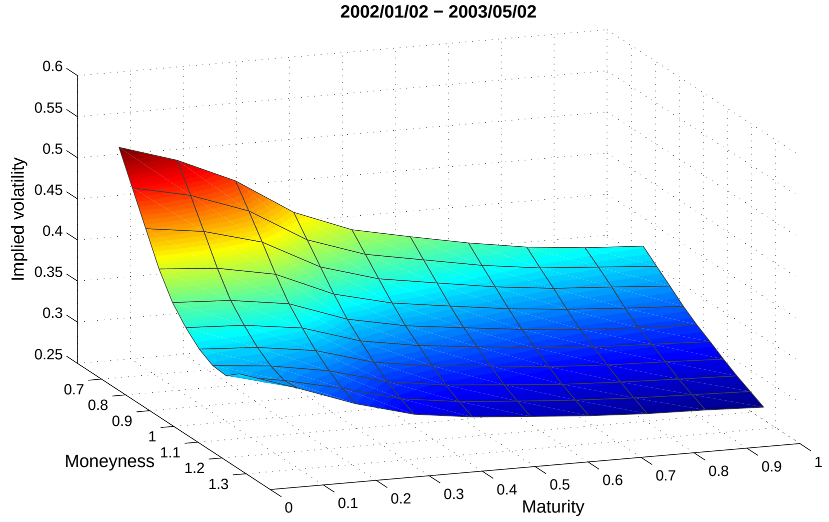


Figure 3.19.: Average DAX IVS for the 1st volatility regime

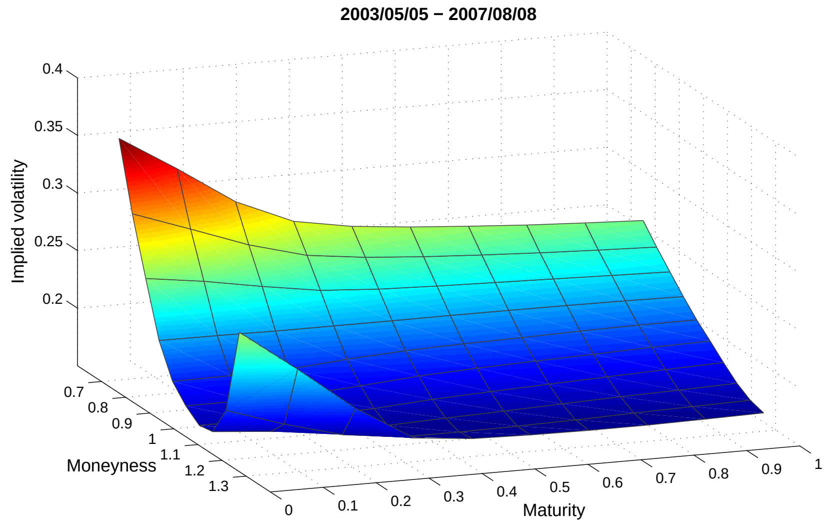


Figure 3.20.: Average DAX IVS for the 2nd volatility regime

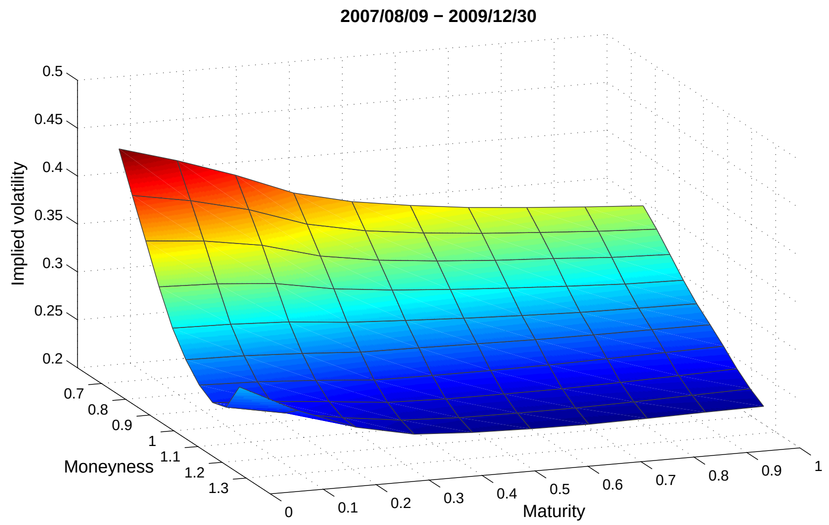
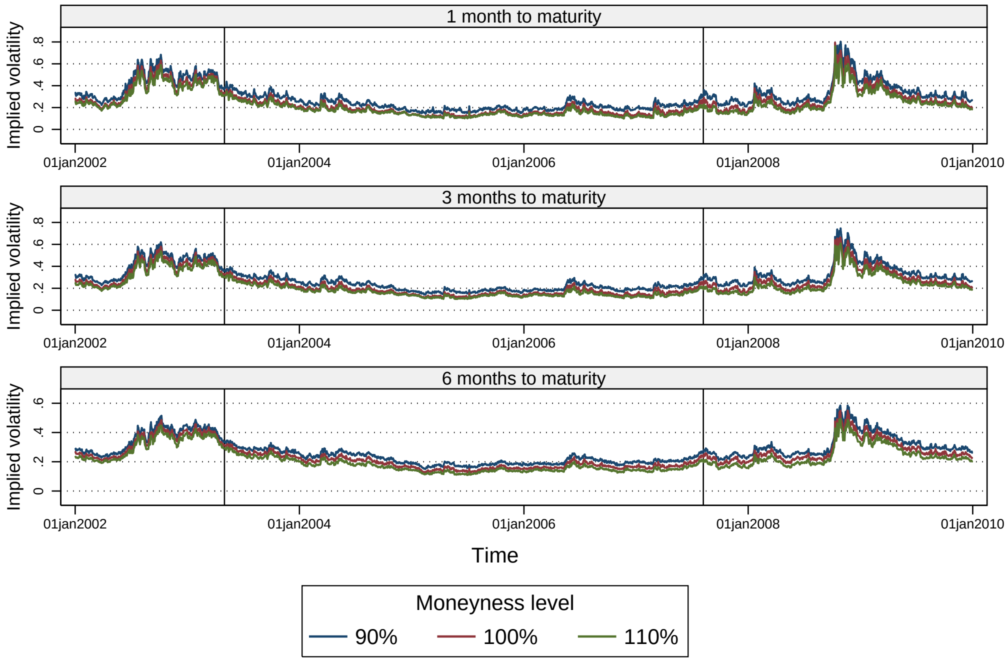


Figure 3.21.: Average DAX IVS for the 3rd volatility regime



Source: EUREX, own calculations.

Note: DAX implied volatilities from 2002 to 2009.

Figure 3.22.: Time series of DAX implied volatilities

3.3.4. Similarities and Differences between DAX and S&P 500 Index Implied Volatility Before and After Stock Market Crashes

In the following, the above results concerning the behaviour of DAX implied volatilities during high volatility market periods are compared to the findings in the literature on the S&P 500 index options market. The aim of the comparison is to examine whether the IVS exhibits certain regularities during turbulent periods (the sample contains two periods which high market fluctuations) that should be accounted when selecting an option pricing model. First, the reversion of implied volatility towards its pre-crash level is discussed. Next, the movements in the implied volatility spreads before and after a stock market crash are considered. Third, the changes in the S&P 500 and DAX volatility term structures during the financial crisis are analysed. Finally, the question of whether implied volatilities before stock market crashes indicated an impending crisis is addressed.

In an analysis of changes in stock market volatility during the financial crisis, Schwert (2011) notes that the implied volatility of S&P index options returned comparatively quickly to normal levels in 2009.[97,98] Based on a long time series of S&P implied volatilities from 1983 to 2010, he argues that the general level of market volatility did not increase after the financial crisis. While a higher implied volatility can be observed after the crisis relative to the pre-crash period, the degree of post-crash implied volatility did not rise relative to the complete sample period. This can be explained by the very low level of market volatility in the period before the financial crisis.[99] Thus, the above finding that DAX implied volatility did not revert back to its pre-crisis level should be qualified by considering not only a longer post-crash period but also a longer pre-crash period.

In a study on the implications of the 1987 crash, Bates (2000) considers the effect of stock market shocks on the magnitude of implied volatility spreads using S&P 500 futures option prices from 1983 to 1993. He illustrates the fundamental change in the implied volatility smile towards the typical post-crash volatility skew based on time series plots of implied volatility spreads for OTM puts and calls.[100,101] Although DAX implied volatility spreads were affected by stock market shocks in 2002 and 2008, a similar, permanent change in the implied volatility spread structure (still) cannot be observed in Figure 3.9.

Next, the changes in the S&P 500 and DAX volatility term structures due to the financial crisis are compared. Schwert (2011) reports that the implied volatilities

---

[97] For instance, during the Great Depression, stock market volatility remained high for several years. See Schwert (2011), p. 796.

[98] Schwert (1990) reports similar results for the stock market crash of 1987.

[99] See Schwert (2011), p. 796.

[100] See Bates (2000), pp. 187-188.

[101] Bates (2000) uses implied volatility spreads that are closely related to the volatility spreads applied in this study. He defines implied volatility spreads as the difference between the implied volatility of an OTM option with a moneyness level of 4% and the implied volatility of an ATM option. Here, moneyness is calculated as $K/F_{SP500} - 1$, where F_{SP500} denotes the S&P 500 futures price. See Bates (2000), p. 187.

of long-term S&P 500 options (with maturities longer than 12 months) were less influenced by the 2008 crisis than those of short-term options (with expiries of less than 6 months). He finds that the implied volatilities of short-term S&P 500 options shifted to a considerably higher level in November 2008. As a consequence, the S&P 500 volatility term structure changed from a relatively flat profile in July 2008 to a downward-sloping shape during the turbulent market period in the fall of 2008.[102] The DAX volatility term structure behaved in a similar fashion. For instance, the right panel of Figure 3.11 depicts a decreasing DAX volatility term structure that was observed in fall 2008. Furthermore, the DAX implied volatility term structure spreads contained in Figure 3.15 imply a downward-sloping term structure during the volatile market period in October/November 2008. With respect to the reversion of the term structure, Schwert (2011) remarks that the S&P 500 volatility term structure returned to its pre-crash shape in April 2010. The positive values of the DAX implied volatility term structure spreads at the end of the sample period (see Figure 3.15) also indicate that the shape of the DAX volatility term structure normalised to its pre-crash pattern by the second half of 2009 (although the level of DAX implied volatility remained above its pre-crisis level).

Bates (1991) also provides a study of the crash of 1987. In the article, Bates (1991) investigates whether market participants expected the crash. He analyses S&P 500 futures options data from 1985 to 1987 and reports evidence that market participants exhibited an increased awareness of downside risk in the year prior to the crash. However, he finds no strong fears of a crash in the two months before the 1987 crash.[103] Similarly, Schwert (1990), using S&P 500 index options, documents that implied volatility did not rise until October 19th, 1987, when the S&P index lost more than 20% of its value.[104] Figure 3.3 indicates that an increased DAX implied volatility level was present from January 2008 to March 2008.[105] However, the DAX implied volatility returned to its previous level at the end of March 2008. In the

[102] Bates (2000) also mentions an inversion of term structure of S&P 500 futures options due to Iraq's invasion of Kuwait in August 1990.

[103] See Bates (1991), pp. 1036-1037.

[104] See Schwert (1990), p. 96.

[105] The decline of the DAX in January 2008 was due to increasing fears of a US recession.

months immediately preceding the financial crisis, no strong increase in volatility could be observed (see Figure 3.3). Therefore, it seems that, similar to the findings of Schwert (1990) regarding the crash of 1987, market participants did not expect the dramatic decline in the DAX generated by the financial crisis.

3.4. Concluding Remarks

In summary, most of the stylised facts concerning DAX implied volatilities that have been documented in the literature for other samples can be confirmed using the options data set employed in this study. However, some important differences were observed, such as the volatility skew for short-term DAX options. Moreover, there are certain similarities with respect to the behaviours of implied volatilities from DAX and S&P 500 index options before and after stock market crashes. Additionally, the analysis demonstrates that the shape of the DAX volatility smile and volatility term structure changed considerably between stable and highly volatile market periods.

Overall, these results demonstrate that the constant volatility assumption of the BS model is violated. Therefore, from the above-described changes in the DAX IVS, it follows that an option pricing model that is sufficiently flexible to allow for these variations is necessary. Therefore, the next Chapter presents some alternative option pricing models and describes their theoretical ability to match the above IVS shapes.

4. Volatility Forecasting Models

In this Chapter, alternative option pricing models are presented that relax some of the BS assumptions. These option pricing models were developed to match the previously described implied volatility patterns. They built on the potential explanations suggested in Chapter 2 for the empirical regularities of the IVS: time-varying volatility of the underlying asset (either deterministically or stochastically) and the occurrence of jumps.[1] As the vast number of option pricing models is overwhelming, the basic concepts of the models are provided in the following Sections. The ability of each model class to reproduce the observed DAX IVS is discussed at the end of each Section. The choice of a particular model from the presented model classes to derive volatility forecasts is explained following the literature review on the forecasting performance of the models in Chapter 5. In addition to the option pricing models, this Chapter also describes time series models that this study applies to forecast DAX volatility.

[1] Additional option pricing models and adjustments to the BS model have been proposed with respect to market frictions. For instance, Leland (1985) suggests an extension of the BS model that allows for transaction costs. However, Constantinides (1996) finds that the volatility smile effect can only be partly attributed to transaction costs. Therefore, this Chapter concentrates on stochastic volatility and jump models.

4.1. Option Pricing Models

4.1.1. Local Volatility Models

Chapter 3.3.2 demonstrated that DAX implied volatilities exhibit certain empirical regularities across maturities, which are called term structure effects. Merton (1973) suggests that these term structure effects can be captured by a time-dependent volatility function.[2] As implied volatility is also affected by the strike price, little additional effort is required to account for this relationship by extending Merton's time-dependent volatility function and allow volatility to also depend on the asset price. Based on this assumption, the stochastic differential equation is given by

$$\frac{dS_t}{S_t} = \mu(S_t, t)dt + \sigma(S_t, t)dz_t. \tag{4.1}$$

The application of this model requires the specification of a parametric form of the volatility function or, alternatively, the application of non-parametric methods.[3] A parametric volatility model was, for example, suggested by Cox and Ross (1976), who were inspired by Fischer Black to develop a general option pricing model in response to the observed negative correlation between stock price changes and volatility changes. Based on the works of John C. Cox and Stephen A. Ross, Fischer Black suggests that this relationship implies the volatility smile.[4] According to Jackwerth and Rubinstein (2001), the negative correlation between stock price and volatility changes is a key contributor to the performance of option pricing models.[5]

More specifically, Cox and Ross (1976) propose a parametric model of the form

$$\frac{dS_t}{S_t} = \mu dt + \sigma S_t^{\beta-1} dz_t \tag{4.2}$$

[2]See Merton (1973), pp. 162-163.
[3]See Detlefsen (2007), p. 54.
[4]See Cox (1996), p. 15.
[5]See Jackwerth and Rubinstein (2001), p. 1.

where $\beta > 0$ is a constant. The model is called a constant elasticity of variance (CEV) model, as the elasticity of variance does not depend on the asset price.[6] For $\beta = 1$, the CEV model corresponds to the classical BS model. In practice, the CEV model exhibits some weaknesses in generating the empirically observed IVS. Detlefsen (2007) argues that this limitation can be explained by the limited number of model parameters.[7,8] To overcome this shortcoming, Beaglehole and Chebanier (2002), Coleman et al. (1999), and Brigo and Mercurio (2001) suggest more flexible models.[9] As a number of parameterisations have been proposed, the best parametric form for the respective application must be selected. Further, except for the mixture diffusions approach of Brigo and Mercurio (2001), these models provide no closed-form solutions. Therefore, option prices have to be calculated based on the BS PDE, in which constant volatility is replaced with the parametric volatility function.[10]

A non-parametric framework to determine the volatility function $\sigma(S_t, t)$ of equation (4.2) was developed by Derman and Kani (1994b), Dupire (1994), and Rubinstein (1994). In this framework, called *local volatility*, it can be shown that a unique *local volatility function* $\sigma(S_t, t)$ that is consistent with a given IVS exists.[11,12] Tree-based algorithms are suggested to extract the local volatility function from current option prices. Although local volatility depends on a stochastic variable, the asset price, it is considered as a deterministic function, as it is uniquely determined by the asset price S_t at time t.[13] For this reason, local volatility models are also called *restricted stochastic volatility models*.[14] Dupire (1994) developed the continuous time theory of the local volatility approach. Derman and Kani (1994b) introduced a discrete-time

[6]See Cox and Ross (1976).
[7]See Detlefsen (2007), p. 55.
[8]See also Rubinstein (1985), who finds that the CEV model does not adequately capture the term structure of US options.
[9]Beaglehole and Chebanier (2002) introduce piecewise quadratic functions, Coleman et al. (1999) apply splines, and Brigo and Mercurio (2001) use mixture diffusions.
[10]See Alexander (2008), p. 248.
[11]See Detlefsen (2007), p. 55.
[12]The graph of the local volatility function is called the local volatility surface (LVS). The term *local* means that the local volatility surface predicts volatility for a certain strike price and time to maturity.
[13]See Hafner (2004), p. 45.
[14]See Brenner (1996), p. 307.

version.[15] To present the general foundation of the approach, the continuous-time version of Dupire (1994) is presented in the following. However, first, local volatility is defined more formally.

In local volatility models, the stochastic evolution of the asset price is driven by a process of the form (4.1). The instantaneous volatility σ is assumed to follow a stochastic process that depends on the asset price S_t. Moreover, the no-arbitrage assumption ensures that a risk-neutral measure exists under which the discounted asset price is a martingale. Finally, the European call option prices $C_t(K,T)$ are given for any strike price K and time to maturity T.[16]

The local variance is defined as the risk-neutral expectation of the squared instantaneous volatility at a future time T conditional on $S_T = K$

$$\sigma^2_{K,T}(S_t, t) = E^{\mathcal{Q}}\{\sigma^2(S_T, T)|S_T = K\} \tag{4.3}$$

and local volatility is given by

$$\sigma_{K,T} = \sqrt{\sigma^2_{K,T}}.^{17} \tag{4.4}$$

In other words, the local volatility approach is based on the assumption that the evolution of instantaneous volatility follows current market expectations, which are captured by the local volatility function. Thus, local volatility can be regarded as the market's consensus perception of instantaneous volatility for a market level K at some future point in time T. In contrast, implied volatility represents the market expectation of the average volatility during the remaining lifetime of the option.[18] Next, Dupire's famous equation that relates unique local volatilities to a cross section of European option prices is derived.

[15]See Gatheral (2006), p. 8.
[16]See Dupire (1994), pp. 18.
[17]See Fengler (2004), pp. 50-51.
[18]See ibid., pp. 49-51.

Assume that the asset price follows the above stochastic process (see equation (4.1)); then, the undiscounted, risk-neutral price of a European call option is

$$C(S_0, K, T) = \int_K^\infty dS_T \phi(S_T, T; S_0)(S_T - K) \tag{4.5}$$

where $\phi(S_T, T; S_0)$ denotes the risk-neutral probability density function of the final asset price at time T.[19] The partial derivatives of (4.5) with respect to K are

$$\frac{\partial C}{\partial K} = -\int_K^\infty dS_T \phi(S_T, T; S_0) \tag{4.6}$$

and

$$\frac{\partial^2 C}{\partial K^2} = \phi(K, T; S_0). \tag{4.7}$$

Further, the density ϕ satisfies the Fokker-Planck equation

$$\frac{1}{2}\frac{\partial^2}{\partial S_T^2}(\sigma^2 S_T^2 \phi) - S\frac{\partial}{\partial S_T}(\mu S_T \phi) = \frac{\partial \phi}{\partial T}\,. \tag{4.8}$$

The first derivative of (4.5) with respect to time is obtained from

$$\frac{\partial C}{\partial T} = \int_K^\infty dS_T \left\{ \frac{\partial}{\partial T}\phi(S_T, T; S_0) \right\} (S_T - K). \tag{4.9}$$

Next, replacing the term $\frac{\partial C}{\partial T}$ in (4.9) with the left hand side (LHS) of (4.8) yields

$$\frac{\partial C}{\partial T} = \int_K^\infty dS_T \left\{ \frac{1}{2}\frac{\partial^2}{\partial S_T^2}(\sigma^2 S_T^2 \phi) - \frac{\partial}{\partial S_T}(\mu S_T \phi) \right\} (S_T - K). \tag{4.10}$$

Partial integration yields

$$\frac{\partial C}{\partial T} = \frac{\sigma^2 K^2}{2}\phi + \int_K^\infty dS_T \mu S_T \phi. \tag{4.11}$$

Substituting (4.6) and (4.7) into (4.11) yields the Dupire equation

$$\frac{\partial C}{\partial T} = \frac{\sigma^2 K^2}{2}\frac{\partial^2 C}{\partial K^2} + \mu(T)\left(-K\frac{\partial C}{\partial K}\right). \tag{4.12}$$

[19] The following derivation of Dupire's equation is taken from Gatheral (2006), pp. 9-11.

If the option price is expressed as a function of the forward price $C(F_T^*, K, T)$, the Dupire function becomes

$$\frac{\partial C}{\partial T} = \frac{\sigma^2 K^2}{2} \frac{\partial^2 C}{\partial K^2} \tag{4.13}$$

where

$$F_T^* = S_0 \exp\left\{\int_0^T dt \mu_t\right\}. \tag{4.14}$$

Rearranging provides the local volatility

$$\sigma^2(K, T, S_0) = \frac{\frac{\partial C}{\partial T}}{\frac{1}{2} K^2 \frac{\partial^2 C}{\partial K^2}}. \tag{4.15}$$

Therefore, based on equation (4.15), local volatilities can be calculated from a set of European option prices.[20] For practical purposes, it should be noted that local volatilities that are directly computed from market option prices using Dupire's equation are highly sensitive to small changes in option prices.[21,22] The computation of local volatilities requires knowledge of the partial derivatives in (4.15). Because discrete sets of options data are typically available, smoothing and extrapolation methods are necessary to obtain continuous data for the calculation of the derivatives.[23] Having received a continuous set of options data, partial derivatives can be approximated using finite differences.[24] The choice of interpolation method should be made carefully, as it can influence the calculation of local volatilities.[25]

While incorporating stochastic volatility and/or jumps into the stochastic process leads to incomplete models, local volatility models impose no additional source of randomness. As a result, the model remains complete and provides a consistent pricing and hedging scheme.[26] Despite the convenient feature of precisely reproducing a set of market option prices, local volatility models exhibit certain drawbacks.

[20] See Gatheral (2006), p. 9.
[21] See Alexander (2008), p. 248.
[22] Avellaneda et al. (1997) suggest a framework for calibrating volatility surfaces via relative-entropy minimisation that is able to produce a smooth volatility surface.
[23] See Chapter 3.1 for a description of smoothing methods.
[24] See Alexander (2008), p. 245.
[25] See Mitra (2009), p. 22.
[26] See ibid., p. 19.

Whereas the one-factor diffusion process inherent to local volatility models permits a complete market model, it does not adequately capture the evolution of the IVS.[27] Empirical studies demonstrate that the assumption of constant local volatilities is incompatible with the observed changes in the IVS. For instance, Alexander and Nogueira (2004a) remark that ignoring other sources of uncertainty in the volatility process leads to unstable local volatility surfaces. In particular, recalibration induces considerable changes in the local volatility surface.[28] Moreover, local volatility models predict that an increasing (decreasing) price of the underlying asset moves the volatility smile to the left (right). This proposed behaviour of the volatility smile contradicts observed market dynamics in which the price of the underlying asset and the volatility smile move in the same direction. This contradiction between the model and the market can induce dynamic hedging problems via unstable delta hedges.[29,30] However, according to Gatheral (2006), the focus of Dupire (1994) and Derman and Kani (1994b) was to develop a model to price exotic options in a manner consistent with the existing volatility smile of vanilla options, rather than a model for volatility dynamics.[31]

Moreover, Fengler (2004) notes that the shape of the observed local volatility surface is typically very "spiky" and counterintuitive.[32] Alexander (2008) argues that the direct calibration of local volatilities to market data can induce spiky local volatility surfaces, as local volatilities are highly sensitive to the input data.[33] Javaheri (2005) notes that this problem can generate arbitrage opportunities and, occasionally, negative variances or probabilities.[34] To avoid this problem, Coleman et al. (1999), Brigo and Mercurio (2001), Beaglehole and Chebanier (2002), and others suggest reconstructing the local volatility surface based on parametric approaches.

[27]See Hafner (2004), pp. 46-47.
[28]See Alexander and Nogueira (2004a), p. 2.
[29]See Hagan et al. (2002), p. 87.
[30]Further, Hagan et al. (2002) show that the delta calculated from local volatility models is incorrect or at least misleading, due to this contradiction. See ibid., p. 84.
[31]See Gatheral (2006), p. 8.
[32]See Fengler (2004), p. 96.
[33]See Alexander (2008), p. 248.
[34]See Javaheri (2005), p. 18.

After a discussion of the general ability of local volatility models to reproduce the IVS, the results of two studies that fit local volatility models to DAX option prices are presented. The first study is by Neumann (1999) and comprises daily DAX option prices from January 1997 to December 1997. He implements the implied binomial tree approach developed by Rubinstein (1994) from the set of local volatility models, the binomial model of Cox et al. (1979), and the BS model to calculate option prices. To evaluate the performance of these models, he computes pricing errors between model and market option prices. He finds that the implied tree approach produces lower pricing errors relative to the BS model and the binomial model.[35] This is not surprising, as the implied tree approach offers greater flexibility than the other two models for matching the observed market prices. Wallmeier (2003) highlights that the essential question, rather, is whether the excellent fit provided by the implied tree approach correctly reflects the possible sample paths under risk-neutrality.[36]

Wallmeier (2003) investigates this question by fitting the implied trinomial trees suggested by Derman et al. (1996) to DAX option prices from 1995 to 2000. Similar to Neumann (1999), he analyses the in-sample fit of an implied tree to the observed market option prices. Further, he compares the future option price at time $t + 7$, which is determined by the implied trinomial tree at t for a certain future asset price level using the market option price at time $t + 7$.[37] His findings agree with those of Skiadopoulos (2001), who summarises some empirical studies providing evidence against the deterministic volatility assumption. In particular, Wallmeier (2003) argues that the empirically observed negative correlation between asset returns and volatility that underlies implied trees cannot completely explain the volatility smile. Thus, the stochastic evolution of the asset price process should incorporate additional factors such as jumps and stochastic volatility.[38] Combining the above-described general features of local volatility models and the findings of

[35] See Neumann (1999), pp. 144-152.
[36] See Wallmeier (2003), pp. 213-214.
[37] See ibid., pp. 226-227.
[38] See ibid., pp. 238-240.

Neumann (1999) and Wallmeier (2003), local volatility models should, in principle, be able to replicate the DAX volatility smiles documented by this study. However, the results of Wallmeier (2003) indicate that local volatility models are problematic when they are used to generate the dynamics of the existing DAX IVS.

Recently, stochastic local volatility models have been developed to overcome these shortcomings.[39] They represent a more general, redefined approach and nested local volatility models.[40] For instance, the Derman and Kani (1998) model allows the local volatility surface to behave stochastically where restrictions ensure the absence of arbitrage.[41] Additional stochastic local volatility models have been developed by Dupire (1996), Britten-Jones and Neuberger (2000), and Alexander and Nogueira (2004b). The following Section represents the Britten-Jones and Neuberger (2000) model.

4.1.2. The Concept of Model-Free Implied Volatility

As the assumption of deterministic volatility is comparatively restrictive, Britten-Jones and Neuberger (2000) extend the local volatility concept developed by Derman and Kani (1994b), Dupire (1994), and Rubinstein (1994). While local volatility models are based on a unique process that allows them to precisely reproduce a complete set of option prices, Britten-Jones and Neuberger (2000) describe a set of continuous processes that are consistent with current option prices. In particular, they derive a condition that must be satisfied by all consistent processes.[42,43] This condition implies the same volatility forecast for all consistent processes. For this reason, the implied volatility calculated based on Britten-Jones and Neuberger (2000) is also called *model-free implied volatility*.[44] Similar to local volatility models, the Britten-

[39]See Fengler (2004), p. 48.
[40]See Skiadopoulos (2001), p. 404.
[41]See Schönbucher (1999), p. 4.
[42]Britten-Jones and Neuberger (2000) note that their approach allows for the application of a variety of stochastic volatility models.
[43]Subsequently, Jiang and Tian (2005) generalised the Britten-Jones and Neuberger (2000) approach to processes with (small) jumps. See Jiang and Tian (2005), p. 1308.
[44]See Britten-Jones and Neuberger (2000), pp. 839-841.

Jones and Neuberger (2000) approach can be used for option pricing. As the focus of this study is to investigate the forecasting performance of implied volatilities, this Section presents the model-free implied volatility forecast and discusses its implementation.[45]

Assume that the underlying asset pays no dividends and the risk-free rate is zero, then the risk-neutral expectation of the integrated variance between T_1 and T_2 is

$$E^Q\left[\int_{T_1}^{T_2}\left(\frac{dS_t}{S_t}\right)^2\right] = 2\int_0^{\infty}\frac{C(T_2,K)-C(T_1,K)}{K^2}dK. \tag{4.16}$$

Based on this equation, the computation of the integrated return variance requires two sets of call option prices with varying K (one set with time to maturity T_1 and the other with T_2). The model-free implied volatility is calculated by taking the square root of (4.16), which yields

$$E^Q\left[\sqrt{\int_{T_1}^{T_2}\left(\frac{dS_t}{S_t}\right)^2}\right] \leq \sqrt{2\int_0^{\infty}\frac{C(T_2,K)-C(T_1,K)}{K^2}dK}. \tag{4.17}$$

From Jensen's inequality, it follows that this is an upward-biased estimator.[46] Next, the implementation of model-free implied volatility as in Jiang and Tian (2005) is described.[47]

The implementation proposed by Jiang and Tian (2005) allows the researcher to relax two assumptions of the Britten-Jones and Neuberger (2000) model. Jiang and Tian (2005) remove the present value of dividend payments from the current stock

[45]Due to the focus of the study, the construction of stochastic volatility processes that are consistent with the initial option prices and can be used to price (exotic) options is not described. A description of this application can be found in Britten-Jones and Neuberger (2000), pp. 848-857.
[46]See ibid., pp. 846-847.
[47]In addition, some citations are taken from Rouah and Vainberg (2007), pp. 322-336.

price to control for dividends. To permit non-zero risk-free rates, they consider call options on forward asset prices. The forward asset price is given by

$$F_t^* = S_t e^{-r(T-t)}.^{48} \tag{4.18}$$

As the call option prices $C(F_t^*, K)$ and $C(S_t, Ke^{rT})$ are identical, the strike price K of the call option in (4.16) is replaced by Ke^{rT} such that

$$E^Q\left[\int_{T_1}^{T_2}\left(\frac{dS_t}{S_t}\right)^2\right] = 2\int_0^\infty \frac{C(T_2, Ke^{rT_2}) - C(T_1, Ke^{rT_1})}{K^2}dK. \tag{4.19}$$

The integrated variance between the current time and time T is obtained from

$$E^Q\left[\int_0^T\left(\frac{dS_t}{S_t}\right)^2\right] = 2\int_0^\infty \frac{C(T, Ke^{rT}) - \max(S_0 - K, 0)}{K^2}dK. \tag{4.20}$$

Applying the trapezoidal rule to approximate the integral yields

$$2\int_0^\infty \frac{C(T, Ke^{rT}) - \max(S_0 - K, 0)}{K^2}dK \approx \sum_{i=1}^m [g(T, K_i) + g(T, K_{i-1})]\Delta K \tag{4.21}$$

where $K_i = K_{\min} + i\Delta K$ and $K_{\min}(K_{\max})$ denote the lowest (highest) available strike price. Further, $\Delta K = (K_{\max} - K_{\min})/m$ represents the difference between two adjacent strike prices where m is equal to the number of strike prices. The function g is defined as $g(T, K_i) = [C(T, K_i e^{rT}) - \max(S_0 - K_i, 0)]/K_i^2$.[49]

Two implementation issues are encountered if model-free volatility is calculated according to (4.21). First, in equation (4.21), model-free volatility is computed by integrating over a complete set of strike prices from zero to infinity. However, in practice, market option prices are only available for a limited strike price interval $[K_{\min}, K_{\max}]$. Thus, by neglecting the tails of the distribution, truncation errors arise. To quantify the extent of these truncation errors, Jiang and Tian (2005)

[48]Here, S_t denotes the stock price after the present value of the dividends to be paid prior to option maturity is removed.

[49]See Jiang and Tian (2005), pp. 1308-1313.

derive upper bounds for right and left truncation errors.[50] Based on a simulation experiment, Rouah and Vainberg (2007) report that the truncation error is smaller than 5%, provided that the strike price range covers the interval $K_{\min} = 0.9K$ and $K_{\max} = 1.1K$.[51] Thus, Taylor et al. (2010) note that the truncation problem is more relevant for stock options, which are traded at fewer strike prices than stock index options.[52] Second, the calculation of the integral requires a continuum of strike prices, which is not available in practice. For instance, the strike price intervals of DAX options are 50, 100, and 200 index points, depending on the remaining lifetime of the option. Jiang and Tian (2005) illustrate the discretisation problem based on the stochastic volatility and random jump model. They report that the discretisation errors of one-month and six-month options are marginal, if $\Delta K \leq 0.35\ SD$, where SD denotes the standard deviation.[53] In addition, Rouah and Vainberg (2007) find that the error is below 4%, if the discrete strike price interval is lower than \$5.00.[54] Despite the particular case of the Heston (1993) model considered by Rouah and Vainberg (2007), Jiang and Tian (2005) present a method to mostly overcome the truncation and discretisation error problems.

They suggest an interpolation-extrapolation method comprising the following steps:

1. In the first step, implied volatilities are calculated based on the observed market options prices via the BS formula.[55,56]

[50] They reported that the truncation errors become very small if the truncation points $K_{\min}$ and $K_{\max}$, which are expressed as multiples of the standard deviation (SD) from the forward price F_0, are more than two 2 SDs away from F_0.

[51] In the experiment, they assume that the volatility of the stock price (which is taken as the reference level for model-free volatility) evolves according to the Heston (1993) model. See Rouah and Vainberg (2007), p. 326.

[52] See Taylor et al. (2010), p. 873.

[53] See Jiang and Tian (2005), p. 1313.

[54] Again, they use the continuous-time variance process of the Heston (1993) model.

[55] Thus, the calculation of model-free volatility is based on all observed option prices, both calls and puts.

[56] While the theoretical concept of Britten-Jones and Neuberger (2000) yields a mostly model-free implied volatility, the implementation developed by Jiang and Tian (2005) uses the BS model to obtain BS-implied volatilities for the smoothing procedure. The question of whether the application of the BS model contradicts the basic concept of model-free volatility is discussed at the end of this Section.

2. To mitigate the truncation problem, the implied volatility surface is extrapolated by setting the implied volatility of options with strike prices below (above) the lowest (highest) available strike price equal to the implied volatility of the lowest (highest) available strike price.[57] Then, a smoothing method is applied to interpolate between the available strike price and thereby increase the discreteness of the strike prices.

3. Call option prices are calculated from the interpolated-extrapolated IVS based on the BS model.

4. The model-free volatility is obtained from (4.21) by plugging the smoothed call option prices into the equation.

The effectiveness of the interpolation-extrapolation method suggested by Jiang and Tian (2005) with respect to mitigating the above mentioned implementation problems is also analysed by Rouah and Vainberg (2007). They demonstrate that the method considerably reduces the truncation and discretisation errors and conclude that, even with a small strike price range, the method delivers a very good approximation result.[58,59] Next, the advantages and disadvantages of the model-free volatility concept are discussed.

The use of BS-implied volatilities to forecast volatility changes has been criticised by many authors, as the BS model is based on constant volatility. Replacing the BS model with a stochastic volatility model to derive implied volatilities resolves this inconsistency (see Section 4.1.3 for a description of stochastic volatility models). However, the application of a stochastic volatility model requires the specification of a stochastic process for the instantaneous volatility. Thus, studies examining options market efficiency based on implied volatilities represent joint tests of market

[57] Jiang and Tian (2005) suggest applying this extrapolation method, as they argue that the approximation error induced by this method is smaller than the truncation error (the truncation error occurs if options with strike prices below (above) $K_{\min}$ ($K_{\max}$) are ignored). See Jiang and Tian (2005), pp. 1316-1318.

[58] See Rouah and Vainberg (2007), p. 332.

[59] The implementation procedure suggested by Jiang and Tian (2005) is, for instance, also used by Muzzioli (2010) and Cheng and Fung (2012).

efficiency and the applied option pricing model. The joint hypothesis problem makes the analysis of options market efficiency difficult, as the misspecification of a given option pricing model can induce systematic errors in option prices and volatility forecasts.[60] Poteshman (2000) argues that these errors can explain why certain studies find that volatility forecasts based on implied volatilities are biased and inefficient.[61] The concept of model-free volatility proposed by Britten-Jones and Neuberger (2000) avoids this problem. In theory, it provides an alternative approach that allows the researcher to test for options market efficiency without relying on a specific option pricing model. However, the use of model-free volatility qualifies this statement, as further assumptions and transformations are necessary.

The above suggested implementation of model-free volatility developed by Jiang and Tian (2005) requires a smoothing method to interpolate between the available strike prices. Principally, the smoothing method can be applied to market option prices and implied volatilities. To avoid numerical difficulties, most studies apply the smoothing method to the implied volatilities.[62] For instance, Martin et al. (2009), Muzzioli (2010), Jiang and Tian (2010), Taylor et al. (2010), and Cheng and Fung (2012) compute model-free volatility by fitting the implied volatility function across strike prices. Smoothing implied volatilities is also widely employed to extract the risk-neutral probability distribution from option prices. Jackwerth (2004) recommends this method, as the magnitude of implied volatilities across strike prices differs less than those for call option prices.[63,64]

However, as the calculation of implied volatilities for the smoothing procedure requires the use of the BS model, the question is whether volatility forecasts based on the Britten-Jones and Neuberger (2000) approach can be still considered model-free. While the use of the BS model to derive volatility forecasts is subject to controversy

[60] See Britten-Jones and Neuberger (2000), pp. 839-840.
[61] See Poteshman (2000), pp. 9-14.
[62] See Jiang and Tian (2005), p. 1315.
[63] See Jackwerth (2004), pp. 20-21.
[64] A comprehensive list of studies on option-implied, risk-neutral distributions that directly apply nonparametric methods to implied volatilities is provided in Jackwerth (2004), p. 22.

in the literature[65], the use of the BS model for the conversion of option prices into implied volatilities (and vice versa) is widely accepted. This transformation adapts the common market practice among option traders of quoting the prices of standard options in terms of implied volatility. Market actors prefer implied volatility to prices when quoting options, as implied volatilities change less in response to movements in the underlying asset price than option prices.[66]

Because, in this context, the BS model resembles a computational tool that delivers a one-to-one mapping between option prices and implied volatilities, Jiang and Tian (2005) argue that it is not necessary for market participants to believe in the assumptions of the BS model.[67] An additional field of research that also considers the application of the BS model as a computational device that does not require accepting the BS assumptions are *stochastic implied volatility models*.[68] Moreover, Hafner (2004) remarks that if the assumptions of the BS model do not hold, then the BS formula is "just a convenient and well-known mapping" function and other bijective transformation functions of the option price are applicable.[69] Because, to my knowledge, the above question regarding the application of the BS model as a simple transformation rule has not been discussed in the literature and is not within the scope of this study, this issue is left for further research.

In addition to the independence of model-free volatility from any option pricing model, the Britten-Jones and Neuberger (2000) approach has further advantages. First, model-free volatility does not require the selection of a particular moneyness level.[70] Second, while BS-implied volatility is based on the information of a single

[65]See, for instance, Campbell et al. (1997), pp. 377-379.
[66]See Schönbucher (1999), p. 2072.
[67]Rebonato (2004), Franke et al. (2004), Fengler (2004), Alexander (2008), and Birru and Figlewski (2012), among others, also note that this practice does not imply that the market actors accept the assumptions of the BS model.
[68]While stochastic volatility models are based on an additional stochastic process for the instantaneous volatility, stochastic implied volatility models directly model the dynamics of the IVS. See the initial work by Schönbucher (1999) and further developments by Brace et al. (2002), Cont and da Fonseca (2002), and Hafner (2004).
[69]See Hafner (2004), p. 37.
[70]Typically, the BS implied volatility of ATM options is used to forecast volatility, as the effects of market frictions on option pricing should be lower for highly liquid ATM options. Further, some studies considering BS implied volatilities at different moneyness levels demonstrate that the BS

option, model-free volatility considers information across all available strike prices. Thus, Bollerslev et al. (2011) argue that using cross-sectional option prices aggregates out pricing errors in individual options.[71] Additionally, by using a broader information set model-free implied volatility should be informationally more efficient than BS-implied volatility.[72]

Because the focus of this study is the forecasting performance of implied volatilities and not information efficiency, the final outcome of the above discussion regarding employing the BS model to obtain implied volatilities does not preclude the use of model-free implied volatility as a volatility forecast.[73] Even if one does not agree with the above-cited papers arguing for the use of the BS model as a transformation rule without accepting the BS assumptions, the use of model-free implied volatility as a volatility forecast can be justified by the larger information set it provides relative to individual BS implied volatilities.[74] In addition, if one completely rejects the use of the BS model in the procedure suggested by Jiang and Tian (2005), it is possible to directly apply the smoothing methods to option prices. An alternative model that is also defined in a stochastic setting is presented in the next Section.[75]

4.1.3. Stochastic Volatility Models

Similar to GARCH models, stochastic volatility models are based on the assumption that volatility follows a stochastic process. Allowing for stochastic volatility introduces a second source of randomness and, as a consequence, market completeness

implied volatility of ATM options provides the most information on future volatility. See, for instance, Fleming (1998) and Christensen and Prabhala (1998).

[71] See Bollerslev et al. (2011), p. 236.

[72] See Jiang and Tian (2005), p. 1337.

[73] However, studies considering the information efficiency of the options market are affected by this discussion due to the above mentioned joint hypothesis problem.

[74] Of course, the argument that model-free implied volatility is not based on a certain option pricing formula is no longer valid.

[75] As the aim of this study is to examine the forecasting performance of implied volatilities, the volatility forecast based on model-free volatility is presented in this Section. The other option pricing models do not provide such a direct volatility forecast. Consequently, the general ability of the Britten-Jones and Neuberger (2000) approach to reproduce the DAX IVS is not considered. A literature review on the forecasting ability of model-free volatility is provided in Chapter 5.

is lost. It follows that a perfect hedge of the option is no longer possible. This hedging problem can be solved by taking a market price for volatility risk into account.[76] Alternatively, other hedging concepts such as super-replication or local risk minimisation can be applied.[77]

The stochastic volatility model of Heston (1993) is currently the most popular stochastic volatility model.[78] The dynamics of the underlying asset price S_t and the variance σ_t^2 are described by

$$dS_t = \mu_t S_t dt + \sigma_t S_t dz_{1,t} \tag{4.22}$$

$$d\sigma_t^2 = \kappa(\overline{\sigma}^2 - \sigma_t^2)dt + \eta\sigma_t dz_{2,t} \tag{4.23}$$

where κ represents the rate of mean reversion of σ_t^2 to its long-run mean $\overline{\sigma}^2$, η is the constant volatility of volatility, and $dz_{1,t}$ and $dz_{2,t}$ are Wiener processes. Thus, the Heston (1993) model assumes that variance is driven by a mean-reverting square-root process that has its own constant volatility. Further, the source of randomness in the volatility process $dz_{2,t}$ is correlated with the randomness of the underlying price process $dz_{1,t}$, the correlation coefficient of which is ρ.[79,80]

To derive Heston's (1993) option valuation equation, the market price of volatility risk must be determined, as the market is incomplete and volatility is, by assumption, not tradable. For this reason, Heston (1993) suggests that the price of volatility risk is proportional to the instantaneous variance in the asset price. This assumption is motivated by asset pricing model developed by Cox et al. (1985), where in equilibrium the consumption process is given by

$$dC_t = \mu_C \sigma_t^2 C_t dt + \sigma_C \sigma_t C_t dz_{3,t}. \tag{4.24}$$

[76] See Rebonato (2004), p. 237.

[77] See Javaheri (2005), p. 27.

[78] See Gatheral (2006), p. 15.

[79] See Heston (1993), pp. 328-329.

[80] The Heston (1993) model was extended by Bates (1996b), Scott (1997), and Pan (2002). Bates (1996b) includes jumps to overcome certain shortcomings of the stochastic volatility model. While Scott (1997) incorporates stochastic interest rates, Pan (2002) adds stochastic dividend payments to the model.

Because the consumption rate C_t is correlated with the asset price, the risk premium (or price of volatility risk) ζ is proportional to the variance, such that $\zeta(S, \sigma^2, t) = \zeta\sigma^2$. The resulting partial differential equation (PDE) is given by

$$\frac{\partial V}{\partial t}+\frac{1}{2}\sigma^2 S^2\frac{\partial^2 V}{\partial S^2}+\rho\eta\sigma^2 S\frac{\partial^2 V}{\partial S\partial\sigma^2}+\frac{1}{2}\eta^2\sigma^2\frac{\partial^2 V}{\partial\sigma^4}+rS\frac{\partial V}{\partial S}-rV = -[\kappa(\overline{\sigma}^2-\sigma^2)-\zeta V]\frac{\partial V}{\partial\sigma^2}.\tag{4.25}$$

[81]

According to Duffie et al. (2000) this equation has a solution of the form

$$C(S, \sigma^2, t) = SP_1 - Ke^{-r(T-t)}P_0 \tag{4.26}$$

where the probabilities P_1 and P_0 must satisfy the above PDE (4.25).[82,83] As the PDE (4.25) must hold for both probabilities P_1 and P_0, the equation (4.26) is substituted into the above PDE. In the following, the PDE is simplified by the introduction of a new variable, which is defined as $x := \ln[S]$. After the substitution, the PDEs can be written as

$$\frac{1}{2}\sigma^2\frac{\partial^2 P_j}{\partial x^2}+\rho\eta\sigma^2\frac{\partial^2 P_j}{\partial x\partial\sigma^2}+\frac{1}{2}\eta\sigma^2\frac{\partial^2 P_j}{\partial\sigma^4}+(r+u_j\sigma^2)\frac{\partial P_j}{\partial x}+(a-b_j\sigma^2)\frac{\partial P_j}{\partial\sigma^2}+\frac{\partial P_j}{\partial t} = 0 \tag{4.27}$$

for $j = 1, 0$ where

$$u_1 = \frac{1}{2}, \quad u_0 = -\frac{1}{2}, \quad a = \kappa\overline{\sigma}^2, \quad b_j = \kappa - j\rho\eta. \tag{4.28}$$

Both PDEs are subject to the terminal condition[84]

$$P_j(x, \sigma^2, T; \ln[K]) = \begin{cases} 1 \text{ if } x \geq \ln[K] \\ 0 \text{ if } x < \ln[K] \end{cases} \tag{4.29}$$

[81] See Heston (1993), p. 329.
[82] See Duffie et al. (2000), pp. 1346-1348.
[83] By analogy with the BS formula, P_1 represents the delta of a European call option and P_0 is equal to the conditional probability that the option is ITM in a risk-neutral world.
[84] See Heston (1993), p. 330.

The stochastic process that drives volatility is a special case of an affine jump-diffusion (AJD) process. Duffie et al. (2000) demonstrate that AJD processes can be solved analytically by calculating an *extended transform*. In the case of the Heston (1993) model, this is a Fourier transform. Thus, based on the Fourier transformation, it can be shown that the desired probabilities P_j are given by the integral of a real-valued function

$$P_j(x, \sigma^2, T; \ln[K]) = \frac{1}{2} + \frac{1}{\pi} \int_0^\infty Re \left[\frac{e^{-i\phi \ln[K]} f_j(x, \sigma^2, T; \phi)}{i\phi} \right] d\phi \tag{4.30}$$

where $f_j(x, \sigma^2, T; \phi)$ denotes a characteristic function.[85]

In the following, the properties of the Heston (1993) model are considered. First, the effect of stochastic volatility on implied volatility is investigated. Next, the implications of the mean-reversion behaviour of volatility are discussed. Finally, the consequences of the volatility of volatility parameter are analysed.

The effect of stochastic volatility on implied volatility depends on the price-volatility correlation ρ. In general, stochastic volatility generates fat tails in the log price distribution. Thus, implied volatilities calculated based on the option prices from the Heston model will exhibit a volatility smile. The shape of the smile pattern is dictated by the correlation coefficient ρ. If volatility and the asset price are not correlated, then the volatility smile is symmetric. A negative (positive) skew occurs when the price-volatility correlation is negative (positive).[86,87]

Next, the effect of mean reversion in volatility on implied volatilities is considered. Mean reversion in volatility is based on the notion that a normal level of volatility exists and volatility tends to revert back towards this level.[88] Therefore, mean reversion governs the behaviour of the volatility term structure. In the Heston (1993) model, mean reversion in volatility is captured by the term $\kappa(\overline{\sigma}^2 - \sigma_t^2)dt$, where $\overline{\sigma}^2$

[85] See Gatheral (2006), pp. 16-19.
[86] See Alexander (2008), p. 271.
[87] The negative price-volatility correlation that is often observed in equity markets can be explained by the *leverage effect*. The leverage effect states that lower stock prices lead to higher firm leverage ratios, which increase stock return volatility. See Hafner (2004), p. 46.
[88] See Engle and Patton (2001), p. 239.

denotes the long-term average of volatility (or normal level) and κ represents the rate of mean reversion.[89] To ensure that the volatility process does not explode, the rate of mean reversion must be positive. Further, higher reversion rates imply that volatility more rapidly converges back to its long-term average level.[90] As Cox et al. (1985) demonstrate that, for positive κ, the variance has a steady-state distribution with mean $\overline{\sigma}^2$, long-term stock returns are asymptotically normally distributed, where the variance per time unit is determined by $\overline{\sigma}^2$.[91]

Finally, the effect of the parameter η, which governs the volatility of volatility, is examined. Simply, if the parameter η is zero, the volatility is deterministic and the stock returns are normally distributed. Heston (1993) demonstrates that a higher volatility of volatility increases the kurtosis and creates heavier tails in the return distribution. This influences the shape of the volatility smile (e.g., a symmetric smile becomes more pronounced for higher η). This means that the Heston (1993) model provides higher far ITM and far OTM option prices than the BS model.[92]

Despite these desirable features, Alexander (2008) notes that the volatility smiles implied by stochastic volatility models for equities and market volatility smiles exhibit different dynamics. For instance, if the underlying stock price changes the model smile, the model and market smiles move in different directions.[93] Another problem associated with the stochastic volatility model is noted by, among others, Bates (1996b) and Das and Sundaram (1999) and refers to the choice of coefficients. They find that unreasonably high parameters are necessary to reproduce the pronounced volatility smiles of short-term options.[94] In addition, as described above, stochastic volatility models are not complete, meaning that the market price of volatility risk must be specified and estimated. This makes the model vulnerable to specification errors.[95] Having discussed the characteristics of stochastic volatility models, the

[89] See Heston (1993), p. 329.
[90] See Alexander (2008), p. 273.
[91] Therefore, the BS model should deliver reasonable results for long-term options. See Heston (1993), p. 335.
[92] See ibid., p. 338.
[93] See Alexander (2008), p. 276.
[94] This problem only occurs for short-term options.
[95] See Hafner (2004), p. 48.

following describes the approaches of Nagel (2001) and Ender (2008), who apply stochastic volatility models to DAX option prices. The intention is to determine whether stochastic volatility models are able to capture the features of the DAX IVS that are outlined in this empirical study (see Chapter 3). First, Nagel's (2001) thesis is considered.

Nagel (2001) compares the option prices calculated from the Heston (1993) model and the BS model with the observed DAX market option prices for the period from February 1992 to December 1994. His out-of-sample results for different moneyness levels and maturities indicate that in most cases the Heston (1993) model exhibits lower pricing errors than the BS model.[96] Thus, the Heston (1993) model provides a better fit for the DAX volatility smile than the BS model.[97]

In a comprehensive study, Ender (2008) extends these findings by investigating the pricing performance of 11 option pricing models, which include the BS model, a stochastic volatility model, a jump-diffusion model, and a stochastic volatility jump-diffusion model.[98] Her sample comprises DAX closing prices for the period from January 2002 to September 2005. Based on an out-of-sample analysis, she reports that the Heston (1993) model can better explain recorded market option prices than the BS model.[99]

The plot of the average DAX volatility smiles in Figure 3.6 demonstrates that DAX options with fewer than 3 months to maturity generally exhibit U-shaped volatility smiles. Further, the DAX volatility smiles of short-term options are more pronounced than those of options with longer maturities. With respect to these patterns, Nagel (2001) finds that the Heston (1993) model is able to capture the U-shaped smile of DAX options, but it is less effective with respect to providing an acceptable fit to the smile for DAX short-term options.[100] Thus, as the patterns of DAX volatility smiles in this study relate to those in Nagel (2001), the application

[96]See Nagel (2001), p. 164.
[97]See ibid., p. 137.
[98]In particular, she employs the Heston (1993), Merton (1976), and Bates (1996b) models.
[99]See Ender (2008), pp. 95-98.
[100]See Nagel (2001), pp. 137-139.

of the Heston (1993) model to fit the existing DAX volatility smile seems suitable for mid-term/long-term options.

Another empirical finding regarding DAX volatility smiles that is considered in Section 3.3.1 is the remarkable time series variation of the smiles. Figure 3.9 and Figure 3.10 show that the levels and skewness of DAX volatility smiles clearly change during the sample period. Note that these variations in the smiles reflect option price changes. Ender (2008) also analyses the ability of the Heston (1993) model to capture such price movements and reports that the model exhibits a better out-of-sample performance than the BS model.[101] Thus, as the data used by Ender (2008) partly overlap the sample of this study, this suggests that the Heston (1993) model is maybe a better choice than the BS model for the present DAX IVS. Whether the introduction of jumps can help to improve these findings is examined in the next Section.

4.1.4. Mixed Jump-Diffusion Models and Pure Jump Models

In addition to stochastic volatility models, *mixed jump-diffusion models* and *pure jump models* (also called *Levy models*) have been suggested to capture the volatility smile and dynamics of the IVS. While mixed jump-diffusion models assume that continuous asset price changes are combined with jumps, pure jump models are based on the assumption that all asset price changes are due to jumps.[102,103] An example of a mixed jump-diffusion model is the Merton (1976) model, which was introduced in Section 2.6.2. As mentioned above, Merton (1976) proposes an extension of the

[101] See Ender (2008), pp. 97-98.

[102] See Hull (2006), p. 562.

[103] As the focus of this Section is to describe the basic principals of jump models, the initial mixed jump-diffusion model developed by Merton (1976) is presented. In addition, a more general jump-diffusion model developed by Bates (1991) is proposed, as it overcomes certain shortcomings of the Merton (1976) model. For a discussion of a pure jump model, e.g., the variance-gamma model, see Hull (2006), pp. 564-566.

BS model in which the asset price follows a continuous diffusion process that is overlaid with jumps. He suggests that the process is given by

$$\frac{dS}{S} = (b - \lambda k)dt + \sigma dz + kdq \tag{4.31}$$

where b denotes the cost-of-carry, λ is the frequency of Poisson events, k represents the average jump size expressed as a percentage of the asset price, and dq is the jump component. Further, he assumes that the two stochastic processes dz and dq are independent.[104]

For the special case that jump sizes are lognormally distributed with variance s^2, a closed-form solution for the price of a European call option is given by

$$\sum_{n=0}^{\infty} \frac{e^{-\lambda' T}(\lambda' T)^n}{n!} C_n(S, K, T, r_n, \sigma_n^2) \tag{4.32}$$

where $\lambda' = \lambda(1+k)$.[105] The BS call option price C_n is determined by

$$\sigma_n^2 = \sigma^2 + \frac{ns^2}{T} \tag{4.33}$$

and

$$r_n = r - \lambda k + \frac{n\gamma}{T} \tag{4.34}$$

where $\gamma = \ln(1+k)$.[106] By imposing the assumption that jump risk is nonsystematic, jump risk can be fully diversified and is thus not priced into the economy. As highlighted in Section 2.6.2, the Merton (1976) model is able to produce pronounced smiles for short-term options that, unfortunately, flatten out too rapidly at longer maturities.[107] The reason is that the effect of jumps decreases at longer maturities, as positive and negative jumps cancel out over long time periods.[108]

[104] See Merton (1976), pp. 132-138.
[105] The parameter k denotes the average jump size, which is expressed as a percentage of the asset price.
[106] See Hull (2006), pp. 563-564.
[107] See Das and Sundaram (1999), p. 213.
[108] See Hafner (2004), p. 49.

Because Merton's simplifying assumption regarding idiosyncratic jump risk is not plausible for index options, Bates (1991) suggests a more general jump-diffusion model that considers the jump component of the asset return as systematic risk. In addition, the Bates (1991) model allows for asymmetric jumps that are consistent with asymmetric volatility skews. Moreover, he assumes that the underlying asset price is, under a risk-neutral probability measure, generated by the following mixed jump-diffusion process:

$$\frac{dS}{S} = (b - \lambda\overline{k})dt + \sigma dz + kdq. \tag{4.35}$$

Further, the logarithm of $1 + k$ is normally distributed

$$\ln(1 + k) \sim N(\gamma - {}^1\!/\!{}_2\, \delta^2, \delta^2) \tag{4.36}$$

and the expected jump size is $E(k) \equiv \overline{k} = e^{\gamma} - 1$.[109]

As Bates (1991) allows for nonsystematic jump risk, some restrictions on preferences and technologies must be imposed.[110] Thus, he assumes that optimally invested wealth follows a jump-diffusion process and consumers have a time-separable power utility function.[111] The risk-neutral jump-diffusion process that describes the asset price changes in the Bates (1991) model is subject to these restrictions.[112]

In the Bates (1991) model, the price of a European call option is given by

$$C = \sum_{n=0}^{\infty} \frac{e^{-\lambda T}(\lambda T)^n}{n!} C_n(S, K, T, r, b_n, \sigma_n) \tag{4.37}$$

with

$$b_n = b - \lambda\overline{k} + \frac{n\overline{\gamma}}{T} \qquad\qquad \sigma_n = \sqrt{\sigma^2 + \delta^2(n/T)} \tag{4.38}$$

[109]See Bates (1991), pp. 1023-1025.
[110]See Bates (1996a), p. 68.
[111]In addition, he makes the assumption that markets are frictionless. See Bates (1991), p. 1024.
[112]See ibid., p. 1024.

where $\overline{\gamma} = \ln(1 + \overline{k})$.[113] The effects of the model parameters on the smile shape are considered below. First, the impact of the jump size is described.

A general feature of jump-diffusion models is that jumps induce fatter tails than would be observed under the normal distribution.[114] As the Bates (1991) model allows for asymmetric jumps, the expected jump size has nonzero mean and can influence the shape of the volatility smile. According to Bates (1991), they are related as follows: if the expected jump size $\overline{k}$ is positive (negative), then the implied distribution is positively (negatively) skewed.[115] As a consequence, negative skewness and excess kurtosis induce an asymmetric volatility smile.[116]

Next, the effect of the jump frequency λ on the smile shape is outlined. In an analysis of the qualitative features of jump-diffusion smiles, Rebonato (2004) demonstrates that an increased jump frequency leads to an (overall) upward shift of implied volatilities and a steeper volatility smile for short-term options. He argues that the overall increase in implied volatilities can be explained by the occurrence of additional jumps, which lead to fatter tails and a higher variance.[117] Despite these attractive features, the application of jump-diffusion models suffers from difficulties in parameter estimation and solution finding for the pricing PDE, as well as the impossibility of perfect hedging.[118] In addition, Das and Sundaram (1999) demonstrate that jump-diffusion models are not consistent with decreasing or U-shaped term structures of ATM forward options.[119] To resolve the question of whether jump models can be employed to replicate the DAX IVS considered in this study, the findings of Detlefsen and Härdle (2007) and Ender (2008) that apply jump models to DAX option prices are considered.

The above-cited study by Ender (2008) also examines the pricing performance of jump models and compares their performance to that of stochastic volatility models

[113] See Haug (2007), pp. 256-257.
[114] See Hafner (2004), p. 49.
[115] See Bates (1991), p. 1024.
[116] See Jarrow (1998), p. 389.
[117] See Rebonato (2004), pp. 494-499.
[118] See Wilmott (1998), p. 333.
[119] See Das and Sundaram (1999), p. 215.

and the BS model for DAX options. While the Merton (1976) model outperforms the BS model, the Merton (1976) model is, in turn, outperformed by the Heston (1993) model.[120] In addition to the Merton (1976) and Heston (1993) models, she also fits the Bates (1996b) model and a variant of Zhu's (2000) modular approach that combines stochastic volatility and jumps. She reports that the use of more complex option pricing models does not result in clearly reduced pricing errors.[121,122]

In a study on calibration risk, Detlefsen and Härdle (2007) fit the Heston (1993) and Bates (1996b) models to DAX options for the period from April 2003 to March 2004. They confirm Ender's (2008) finding that the Heston (1993) and Bates (1996b) models provide a good fit for the DAX IVS.[123,124] As neither of these studies identifies clear differences between the performance of the Heston (1993) and Bates (1996b) models, the use of a simpler model (here: the Heston (1993) model), which is less prone to overfitting, is advisable.

Because the samples used in Ender (2008) and the present overlap, Ender's (2008) findings are relevant for the above question regarding fitting the DAX IVS analysed in this study using jump models. Thus, drawing on the evidence presented by Ender (2008), the Merton (1976) model should better support the characteristics of the existing DAX IVS than the BS model. However, comparing jump models with stochastic volatility models, the above results suggest the use of the Heston (1993) model to fit the DAX IVS considered in this study.

In addition to these considerations, the abovementioned inconsistency of jump-diffusion models with certain patterns of the volatility term structure, which is documented by Das and Sundaram (1999), is essential for reproducing the DAX IVS. Figure 3.12 illustrates that the average term structure is decreasing for short-

[120] These results are based on an out-of-sample analysis.
[121] See Ender (2008), pp. 95-98.
[122] With respect to the comparison of jump models and the BS model, she finds that the Merton (1976) and Bates (1996b) models generate lower pricing errors than the BS model. See Ender (2008), pp. 97-98.
[123] See Detlefsen and Härdle (2007), p. 52.
[124] Detlefsen and Härdle (2007) investigate the pricing performance of the models using in-sample tests. See ibid., p. 51.

term DAX ATM options. Further, the term structure spreads presented in Figure 3.15 indicate that the slope of the DAX term structure changes during the sample period. For this reason, jump-diffusion models do not seem to be a reasonable choice to fit the DAX IVS considered in this study. Besides forecasting financial volatility based on implied volatility, time series models can be used to predict volatility. The following Section presents a brief introduction into some time series models that are applied in this study to forecast DAX volatility.

4.2. Time Series Models for Forecasting Volatility

In their comprehensive review on forecasting financial volatility, Poon and Granger (2003) present the most commonly used time series volatility models. They classify time series volatility models into three groups. The first group is called *Predictions Based on Past Standard Deviations* and covers the *Historical Average*, the *Moving Average*, the *Exponential Smoothing* and the *Exponentially Weighted Moving Average (EWMA)* methods.[125] Further, they include the class of *Autoregressive (AR)* models such as the *Autoregressive Integrated Moving Average (ARIMA)*, *ARFIMA*, and *Threshold Autoregressive (TAR)* models, which relate volatility to its lagged values and past error terms, in this group. In addition, the *HAR* model developed by Corsi (2009) also belongs to this group.[126] The second group of time series volatility models comprises the ell-known family of *Autoregressive Conditional Heteroskedasticity (ARCH)* and *GARCH* models, which focus on conditional variance. The last group contains *stochastic volatility (SV)* models that describe instantaneous

[125] Poon and Granger (2003) do not cover nonparametric methods to predict volatility, as some studies find that these methods have low forecasting power. They refer to Pagan and Schwert (1990) and West and Cho (1995).

[126] Poon and Granger (2003) do not mention the HAR model, as it had yet to be published. However, the models belonging to this group follow its model structure.

volatility using an own stochastic process. The GARCH model family, the ARFIMA model, and the HAR model are presented in the following.[127,128]

4.2.1. GARCH Models

Financial time series exhibit certain stylised facts that are well-documented in the literature. An important finding is that the distribution of financial returns exhibits fatter tails than the normal distribution. The fatter tails are induced by an increased number of outliers and volatility clustering. The volatility clustering effect describes the tendency of financial volatility to cluster. This means that a large (small) price change is followed by another large (small) price change. Engle (1982) and Bollerslev (1986) developed a volatility class of models, the ARCH and GARCH models, to capture these stylised facts. As many excellent introductions to GARCH models exist[129], the following Section concentrates on providing an introduction to the basic model structure, the stationarity condition, and the forecast equation. First, the standard GARCH model is presented.

The GARCH Model

Suppose that the asset return a_t is explained by

$$a_t = \mu + \varepsilon_t \tag{4.39}$$

$$\varepsilon_t = \sqrt{h_t} u_t \tag{4.40}$$

[127] See Bauwens et al. (2012) as well as Xiao and Aydemir (2011) for a comprehensive description of recent volatility models.

[128] As the estimation of SV models is more complex than for GARCH models (see Martino et al. (2011), p. 487), this study does not use SV models. Moreover, more simplistic approaches (e.g., the EWMA method) are not employed, as some of them can be attributed to more sophisticated models (in case of the EWMA method, this is the so-called *Integrated GARCH*(1,1) model), and their DAX forecasting performance is documented by other studies.

[129] See, for instance, Franke et al. (2004), Teräsvirta (2009), and Francq and Zakoian (2010).

where $u_t \sim D(0,1)$ denotes a white noise process.[130,131] A process (ε_t) is called a GARCH(p,q) process if it satisfies

1. $E(\varepsilon_t|I_{t-1}) = 0, \quad t \in \mathbb{Z}$
2. $Var(\varepsilon_t|I_{t-1}) = h_t = \omega + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j h_{t-j}, \quad t \in \mathbb{Z}$

where I_{t-1} denotes the information set that contains all information up to time $t-1$.[132] To ensure that the conditional variance is always positive, the sufficient, but not necessary, conditions are $\omega > 0$, $\alpha_i \geq 0$ for $i = 1, \ldots, q$ and $\beta_j \geq 0$ for $j = 1, \ldots, p$. Although the GARCH model is parsimoniously parameterised, it allows the researcher to capture persistent volatility clusters.[133]

The covariance stationarity of a GARCH(p,q) process requires that

$$\sum_{i=1}^{q} \alpha_i + \sum_{j=1}^{p} \beta_j < 1. \tag{4.41}$$

In this case the unconditional variance is given by

$$\sigma^2 = \frac{\omega}{1 - \sum_{i=1}^{q} \alpha_i - \sum_{j=1}^{p} \beta_j}.^{134} \tag{4.42}$$

In empirical applications, a GARCH (1,1) is often sufficient to describe the data. The conditional variance of a GARCH(1,1) has the form

$$h_t = \omega + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1} \tag{4.43}$$

with $\omega > 0$, $\alpha_1 > 0$, and $\beta_1 > 0$. The process exhibits weak-stationarity for $\alpha_1 + \beta_1 < 1$. The kurtosis of a GARCH(1,1) process exists if $3\alpha_1^2 + 2\alpha_1\beta_1 + \beta_1^2 < 1$ and is given by

$$K[\varepsilon_t] = 3 + \frac{6\alpha_1^2}{1 - \beta_1^2 - 2\alpha_1\beta_1 - 3\alpha_1^2}. \tag{4.44}$$

130 See Poon (2005), p. 37.
131 It is frequently assumed that the distribution D is normal.
132 See Francq and Zakoian (2010), p. 19.
133 See Kirchgässner and Wolters (2007), pp. 252-254.
134 See Poon (2005), p. 38.

It is always greater than three because $\alpha_1 > 0$. This implies that the distribution of ε_t is leptokurtic.[135]

Further, it can be shown that the equation for the GARCH(1,1) model can be written as

$$h_t = \frac{\omega}{1-\beta_1} + \alpha_1 \sum_{j=1}^{\infty} \beta_1^{j-1} \varepsilon_{t-j}^2. \tag{4.45}$$

This transformation demonstrates that the GARCH(1,1) has an ARCH(∞) representation that is characterised by geometrically declining weights.[136]

Finally, this Section presents the construction of (multi-)period volatility forecasts based on the GARCH(1,1). The one-step-ahead volatility forecast based on the information set I_t is given by

$$\widehat{h}_{t+1|t} = E[\varepsilon_{t+1}^2 | I_t] = \omega + \alpha_1 \varepsilon_t^2 + \beta_1 h_t. \tag{4.46}$$

Similarly, the two-step forecast at time t is

$$\widehat{h}_{t+2|t} = \omega + \alpha_1 \varepsilon_{t+1}^2 + \beta_1 h_{t+1}. \tag{4.47}$$

Based on $E[\varepsilon_{t+1}^2 | I_t] = h_{t+1}$[137], which follows from (4.40), the forecast of h_{t+2} can be expressed as

$$\widehat{h}_{t+2|t} = \omega + (\alpha_1 + \beta_1) h_{t+1}. \tag{4.48}$$

Repeated substitution for the τ-step-ahead forecast yields

$$\widehat{h}_{t+\tau|t} = \omega \frac{1-(\alpha_1+\beta_1)^{\tau-1}}{1-\alpha_1-\beta_1} + (\alpha_1+\beta_1)^{\tau-1} h_{t+1|t}. \tag{4.49}$$

For large τ, the forecast of the conditional variance converges towards its unconditional variance.[138]

---

[135] See Franke et al. (2004), pp. 218-220.

[136] See Francq and Zakoian (2010), p. 42.

[137] Note that $E[\varepsilon_t^2] = h_t$.

[138] See Kirchgässner and Wolters (2007), pp. 255-256.

The Exponential GARCH model

Standard GARCH models are based on the assumption that positive and negative shocks have the same effect on conditional volatility. However, empirical studies typically find that bad news (negative shocks) tends to increase volatility to a greater extent than good news (positive shocks).[139] To extend the GARCH model to represent such asymmetric effects, Nelson (1991) suggests the *Exponential GARCH* (EGARCH) model.[140] The EGARCH(p,q) models specifies a logarithmic formulation for the conditional variance, which can be written as

$$\ln h_t = \omega + \sum_{j=1}^{p} \beta_j \ln h_{t-j} + \sum_{k=1}^{q} \left[\theta_k \epsilon_{t-k} + \gamma_k \left(|\epsilon_{t-k}| - \sqrt{2/\pi} \right) \right] \tag{4.50}$$

where $\epsilon_t = \varepsilon_t / \sqrt{h_t}$.[141] Because this formulation ensures that the conditional variance is always positive, no further non-negativity conditions are necessary. The ARCH effect in equation (4.50) is produced by the term $\theta_k \epsilon_{t-k}$, and the asymmetric effect is captured by $\gamma_k \left(|\epsilon_{t-k}| - \sqrt{2/\pi} \right)$. The covariance stationarity of the process is guaranteed by $\sum_{j=1}^{q} \beta_j < 1$.

Due to the logarithmic formulation of the EGARCH model, the calculation of the forecasts is more sophisticated than those of the standard GARCH model. For instance, according to Tsay (2005), the one-step-ahead volatility forecast based on the EGARCH(1,0) is given by

$$\widehat{h}_{t+1} = h_t^{2\alpha 1} e^{(1-\alpha_1)\omega} e^{g(\epsilon)} \tag{4.51}$$

$$g(\epsilon) = \theta \epsilon_{t-1} + \gamma \left(|\epsilon_{t-1}| - \sqrt{2/\pi} \right). \tag{4.52}$$

[139] Black (1976) was the first to describe this effect. He attributes it to the following mechanism: bad news tends to reduce the price of an asset, which implies a higher debt-to-equity ratio and thus higher volatility. Therefore, this effect is also called the *Leverage Effect*.

[140] In addition to the EGARCH model, Glosten et al. (1993) and Zakoian (1994), among others, have also proposed asymmetric models. As numerous GARCH specifications exist, it is necessary to select the relevant GARCH models. An explanation of why this Section of this study employs particular GARCH models is provided in Section 6.4.

[141] In addition, it is assumed that $E(\epsilon) \sim N(0, 1)$.

The multi-step forecast has the form

$$\widehat{h}_{t+\tau} = h_t^{2\alpha 1}(\tau - 1)e^{\eta}[e^{0.5(\theta-\gamma)^2}\Phi(\theta+\gamma) + e^{0.5(\theta-\gamma)^2}\Phi(\theta-\gamma)] \tag{4.53}$$

where

$$\eta = (1-\alpha_1)\omega - \gamma\sqrt{2/\pi} \tag{4.54}$$

and Φ denotes the standard normal cumulative distribution.

4.2.2. Long Memory Models

Empirical studies have demonstrated that a number of financial time series, such as nominal and real interest rates, real exchange rates, and realised volatility, are often highly persistent.[142,143] Principally, time series can exhibit persistence in the first and higher order moments. While GARCH models are typically able to capture persistence in the volatility of a time series, stationary ARMA models typically cannot adequately capture the persistence in the first moment.[144] Further, the common practice of taking the first differences to analyse a stationary time series may be overly extreme and misleading.[145] For this reason, long memory models have been suggested to describe (highly) persistent financial time series. In the following, the ARFIMA models developed by Granger and Joyeux (1980) and Hosking (1981) are presented.

ARFIMA models

Long memory time series are characterised by a very slowly decaying autocorrelation function. To address this long-range dependence using an ARMA process, a

[142] See, for instance, Andersen et al. (2001a) and Baillie et al. (1996).
[143] Here, realised volatility is defined as the square root of the sum of squared intraday returns. See Section 6.6.1 for a formal definition of realised volatility and an introduction to the concept of realised volatility.
[144] See Zivot and Wang (2008), p. 271.
[145] See Box et al. (2008), p. 429.

large number of autoregressive terms are necessary.[146] Granger and Joyeux (1980) and Hosking (1981) demonstrate that a long memory process can be described by fractionally integrated processes in a more parsimonious way. In ARFIMA models, the integration order is no longer restricted to integer values. They allow researchers to model persistence or long memory using a fractionally integrated I(d) process in which the fractional integration parameter d is restricted to $0 < d < 1$.[147] Thus, ARFIMA processes close the gap between short- and long-run memory models by allowing for short-run and long-run dependencies.[148]

A time series y_t is called an ARFIMA(p, d, q) process if

$$A(L)(1-L)^d(y_t - \mu) = B(L)\varepsilon_t, \quad \varepsilon_t \sim iid(0, \sigma^2) \tag{4.55}$$

where $A(L) = 1 - \alpha_1 L - \ldots - \alpha_p L^p$ and $B(L) = 1 - \beta_1 L - \ldots - \beta_q L^q$ are lag polynomials with roots outside the unit circle.[149,150] The characteristics of the time series depend on the value of d. For $0 < d < 0.5$, the process is covariance stationary and has long memory. When $0.5 < d < 1$ the process still captures long memory effects, but the series is no longer stationary.[151] The long memory parameter is typically estimated using the log periodogram estimator of Geweke and Porter-Hudak (1983).[152] The following addresses producing forecasts using an ARFIMA(p, d, q) model with external regressors.[153]

[146]See Zivot and Wang (2008), pp. 272-273.
[147]See Granger and Joyeux (1980), pp. 15-16.
[148]See Franke et al. (2011), p. 357.
[149]See Granger and Joyeux (1980), pp. 16-21.
[150]While Li (2002) fits an ARFIMA model to the linear realised volatility series $y_t = \sigma_t^2$, Martens and Zein (2004) and Pong et al. (2004) use an ARFIMA model for the log-realised volatility series $y_t = ln(\sigma_t^2)$.
[151]See Franke et al. (2011) who present a summary of time series long memory characteristics for different fractional integration parameters d. See Franke et al. (2011), p. 347.
[152]Alternatively, Robinson (1995a) suggests a Gaussian semiparametric estimator.
[153]This more general model comprises the ARFIMA model given by (4.55). For instance, Martens et al. (2009) use an extended ARFIMA model that allows for a gradual level shift, day-of-the-week effects, and nonlinear effects of lagged returns to fit realised S&P 500 volatilities. They report that the extended ARFIMA model provides better one-day forecasts than the linear ARFIMA model. See Martens et al. (2009), pp. 291.

An ARFIMA(p, d, q) model with k regressors is given by

$$A(L)(1-L)^d(w_t) = B(L)\varepsilon_t, \quad \varepsilon_t \sim iid(0, \sigma^2) \tag{4.56}$$

where $w_t = y_t - \mathbf{x}_t\beta$. The best linear prediction of $w_{t+\tau}$ is

$$\widehat{w}_{t+\tau|t} = (\gamma_{t-1+\tau} \ldots \gamma_\tau)(\mathbf{\Sigma}_t)^{-1}\mathbf{w} = \mathbf{q}'_\tau\mathbf{w} \tag{4.57}$$

where $\gamma_\tau, \tau = 1, \ldots, t$ denotes the autocovariance function of the ARFIMA(p, d, q) process. This procedure is equivalent to a regression of $w_{t+\tau}$ on $\mathbf{w}$, where $(\mathbf{\Sigma})_t^{-1}\mathbf{w}$ can be determined by the Durbin-Levinson algorithm.[154,155]

The HAR Model

In addition to ARFIMA models, Corsi (2009) suggests a simple AR-type model for realised volatility that is also able to reproduce long memory and fat tails. Due to its additive structure, which consists of the sum of three volatility components, it is called a heterogeneous autoregressive (HAR) model.[156] The model is motivated by the Heterogeneous Market Hypothesis advanced by Müller et al. (1995), which states that market participants differ with respect to risk aversion, institutional constraints, degrees of information, prior beliefs, transaction costs and other characteristics.[157] This heterogeneity implies "that different market agent types or *components* perceive, react to, and cause different types of *volatility*"[158]. In particular, as heterogeneous market agents are sensitive to different time horizons, they are active in the market at different frequencies (e.g., fund managers versus day traders) and cause different volatility components.[159] Further, the HAR model accounts for the asymmetric reactions of agents with different investment horizons to volatility

[154] See Doornik and Ooms (2006), p. 6.
[155] For a more comprehensive description, see Beran (1994).
[156] Empirical applications of the HAR model and its extensions are provided by Andersen et al. (2007), Martens et al. (2009), and Busch et al. (2011).
[157] See Müller et al. (1995), p. 12.
[158] See Müller et al. (1997), p. 214.
[159] See Müller et al. (1995), p. 12.

changes. While short-term traders also consider long-term volatility, investors with longer holding periods will not generally revise their trading strategies when the level of short-term volatility changes.[160,161] To capture these dependencies, Corsi (2009) suggests a cascade model that specifies current volatility as the sum of past volatility components over different horizons.

The HAR model is defined as

$$RV_t = \alpha + \beta^d RV_{t-1} + \beta^w RV_{t-5:t-1} + \beta^m RV_{t-22:t-1} + \varepsilon_t \tag{4.58}$$

where $RV_{t+1-k:t} = \frac{1}{k}\sum_{j=1}^{k} RV_{t-j}$ denotes multiperiod volatility components and $\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$.[162] Several extensions have been proposed to capture other stylised facts such as jumps and the leverage effect.[163] Corsi (2009) demonstrates that while the HAR model does not formally belong to the class of long memory models, the model is able to reproduce hyperbolic decaying sample autocorrelations.[164,165]

In practice, the forecasting performance of the HAR model is typically comparable to that of ARFIMA models. This is an important result, as the HAR model has a simple structure.[166] Wang et al. (2013) consider autoregressive models to forecast long memory processes that are subject to structural breaks and provide an explanation for this remarkable performance. They argue that the HAR model

[160] See Aït-Sahalia and Mancini (2008), p. 23.

[161] Short-term traders consider changes in long-term volatility, as these changes reflect potential effects on expected trends and riskiness (see Aït-Sahalia and Mancini (2008), p. 23). In contrast, short-term volatility is irrelevant for market agents with a long-term investment horizon. See Louzis et al. (2012), p. 3535.

[162] Alternatively, the model can be formulated for log-volatility. Corsi et al. (2008) find that the log-volatility specification of the HAR model yields better forecasting performance. See Corsi et al. (2008), pp. 69-74.

[163] See, for instance, Andersen et al. (2007), who separate the jump and continuous volatility components.

[164] See Corsi (2009), pp. 176-186.

[165] This behaviour is not surprising, as Granger (1980) demonstrates that aggregating a large number of AR(1) processes where the autoregressive parameters are taken from a particular Beta distribution on $(0, 1)$ yields a long memory process. See Granger (1980), pp. 230-234 and Granger and Ding (1996), p. 71.

[166] See Hansen and Lunde (2011), p. 542.

avoids inaccurate parameter estimation and spurious breaks.[167] However, despite their successful forecasting performance, Gregoriou (2009) argues that HAR models are difficult to further justify.[168]

Finally, the HAR forecasting equation for log-realised volatility is given by

$$\ln RV_{t+\tau|t} = \alpha + \beta^d \ln RV_t + \beta^w \ln RV_{t-4:t} + \beta^m \ln RV_{t-21:t} + \varepsilon_t \tag{4.59}$$

where $RV_{t+\tau|t}$ denotes the average log-realised volatility between t and $t+\tau$ for $\tau \geq 1$.

[167] They attribute the weak forecasting performance of ARFIMA models relative to AR models in finite samples to estimation problems. In particular, they note that if the fractional integration parameter is close to 0.5, d is difficult to estimate (see Aït-Sahalia and Mancini (2008), p. 176). Moreover, structural break tests often experience difficulty in distinguishing between actual and spurious breaks when the data generating process is fractionally integrated. See Kuan and Hsu (1998), pp. 705-706.

[168] See Gregoriou (2009), p. 427.

5. Forecasting Performance of Volatility Models: A Literature Review

This Chapter presents a literature review of empirical studies comparing the volatility forecasting performance of implied volatility and time series models. Because most early studies use encompassing regressions to evaluate volatility forecasts, the first Section introduces this evaluation method. The second Section reviews, with one exception, selected papers on predicting US stock market volatility, as these articles contain broad and intensive discussions of the US stock market while there is no analogous discussion of the German stock market.[1] The findings of these papers concerning the applied forecasting models and evaluation methods provide useful information for the empirical analysis performed in this thesis. While chapters 5.2.1 to 5.2.5 present studies using encompassing regressions to evaluate volatility forecasts, the last Section of Chapter 5.2 discusses the forecasting performance of volatility models based on statistical loss functions.[2] The following Section introduces empirical studies on the predictive ability of implied volatility and time series models for German stock market volatility. Based on the results of these studies, the model

[1] The exception is the study by Li (2002) that considers foreign exchange rates. The results of this study are presented in Chapter 5.2, because it is the first study comparing the forecasting ability of implied volatility and long memory models based on realised volatility.

[2] As a alternative to encompassing regressions, relative model performance can also be measured by ratios or differences in mean, mean-squared, or mean-absolute prediction errors. For a description of this forecast evaluation technique see Chapter 6.6.2.

characteristics described in Chapter 4, and the analysis of DAX (implied) volatility presented in Chapter 3, the final Section explains the choice of the volatility prediction models used in this study to forecast DAX volatility. Because the focus of this thesis is on volatility prediction, the following discussion generally presents out-of-sample results.

5.1. Volatility Forecast Evaluation Based on Encompassing Regressions

5.1.1. The Definition of Information Efficiency

Numerous papers apply encompassing regressions to investigate the correlation between predicted and realised volatility (also called ex-post volatility).[3] Most studies using encompassing regressions to examine the forecasting performance of implied volatilities and historical volatility consider the so-called informational efficiency of the options market. According to Malkiel (1992), an information efficient market is defined as follows:

> A capital market is said to be efficient if it fully and correctly reflects all relevant information in determining security prices. Formally, the market is said to be efficient with respect to some information set, ϕ, if security prices would be unaffected by revealing that information to all participants. Moreover, efficiency with respect to an information set, ϕ, implies that it is impossible to make economic profits by trading on the basis of ϕ.[4]

[3]See West (2006), p. 101

[4]See Malkiel (1992). p. 739.

Various papers investigate the information efficiency of options markets.[5] To derive a testable market efficiency hypothesis, empirical studies typically characterise the informational efficiency of the options market as follows:

> If markets are efficient and the option pricing model is correct, the implied volatility calculated from option prices should represent the market's best forecast of the underlying asset's future volatility over the remaining life of the option. As such, it should be both unbiased and informationally efficient — that is, it should correctly impound all available information, including the asset's price history.[6]

Therefore, encompassing regressions are often applied to investigate the information efficiency of options markets. Specifically, such papers examine whether implied volatilities reflect (all) relevant available information to forecast stock market volatility. To test this hypothesis, the realised volatility of the underlying asset is regressed on implied volatility and historical volatility (in early studies, typically the lagged standard deviation, while recent volatility forecasts are based on more sophisticated time series models). The regression tests are presented below.

5.1.2. Encompassing Regressions

Most studies investigating the forecasting performance of implied volatility and time series models are based on estimations of the following regressions:

$$\sigma_t^r = \alpha_0 + \alpha_1 \sigma_t^{iv} + \varepsilon_t, \tag{5.1}$$

$$\sigma_t^r = \beta_0 + \beta_1 \sigma_t^{ts} + \varepsilon_t, \tag{5.2}$$

and

$$\sigma_t^r = \gamma_0 + \gamma_1 \sigma_t^{iv} + \gamma_2 \sigma_t^{ts} + \varepsilon_t \tag{5.3}$$

[5]See, for example, the studies cited in this Chapter.
[6]See Ederington and Guan (2002), p. 29.

where σ_t^r denotes the realised volatility of the underlying asset, σ_t^{iv} is the implied volatility, and σ_t^{ts} represents a volatility forecast (also referred to as historical volatility) from a time series model.

Because the volatility of the underlying asset returns is unobservable, a volatility proxy using realised volatility measures is needed. Up to the availability of intraday returns, ex-post volatility has typically been measured using the sample standard deviation of the underlying asset returns over a fixed horizon or the remaining life of the option. In more recent studies, high-frequency returns have become an increasingly common method to estimate ex-post volatility.[7]

The sample standard deviation of past returns has also been employed to produce volatility forecasts. Recently, more sophisticated time series models, e.g., GARCH models, are applied to generate volatility predictions. Similarly, more recent studies employ broader set of option pricing models to calculate implied volatility. While initially the BS option pricing model was widely used to compute implied volatility, the most recent studies also apply alternative option pricing models, e.g., model-free implied volatility or stochastic volatility models.

Various hypotheses are tested using the regression equations presented above. In particular, the univariate regressions can be used to examine the following hypotheses. First, if implied or historical volatility contains information about future volatility, then the estimated slope coefficients α_1 and β_1, respectively, should be significantly different from zero. Second, the unbiasedness of the volatility forecasts is verified by $\alpha_0 = 0$ and $\alpha_1 = 1$ or $\beta_0 = 0$ and $\beta_1 = 1$. Further, the relative information content of the volatility forecasts is analysed by comparing the R^2 values of the univariate regression equations.

Encompassing regressions allow the researcher to investigate the following three hypotheses. The first hypothesis states that implied volatility subsumes all information that is contained in historical volatility regarding future volatility and is examined

[7] See Chapter 6.6.1 for a discussion of volatility proxies.

by testing whether $\gamma_2 = 0$. The second hypothesis further requires that the coefficient γ_1 is equal to one. The most restrictive hypothesis states that implied volatility is unbiased and efficient and is tested by verifying that $\gamma_0 = 0$, $\gamma_1 = 1$, and $\gamma_2 = 0$. The empirical studies presented in the next Section consider forecasting approaches for US stock market volatility and are based on this forecast evaluation technique.

5.2. Empirical Studies Forecasting US Stock Market Volatility

5.2.1. The Initial Debate over the Predictive Ability of Implied Volatility

The first studies concerning the predictive power of implied volatility for US stock market volatility are presented by Latané and Rendleman (1976), Schmalensee and Trippi (1978), Chiras and Manaster (1978), and Beckers (1981). These papers provide evidence that implied volatility exhibits better predictive ability than the lagged standard deviation.[8,9]

Day and Lewis (1992) extend these findings and compare weekly volatility forecasts based on BS ATM implied volatility, the GARCH model, and the EGARCH model. They consider S&P 100 index call options from March 1983 to December 1989. The forecasts are evaluated by estimating the above regressions where ex-post volatility is measured by two volatility proxies.[10] Drawing on the above univariate regressions they demonstrate that the estimated coefficients of the volatility forecasts, α_1 and β_1, are, with one exception, significantly different from zero.[11] The estimation results show that the intercepts of the volatility forecasts based on the GARCH model,

[8] See Mayhew (1995) for a brief review of the articles.

[9] The results provided by these early studies must be interpreted with care, as they suffer from several shortcomings. See, e.g., Canina and Figlewski (1993) and Ederington and Guan (2002).

[10] They use the square of the weekly return and the variance of the week's daily returns.

[11] Additionally, they report that the R^2 values for the regressions are rather low.

β_0, are not significantly different from zero and that the β_1 coefficients are not different from one. This implies that the GARCH forecasts of conditional volatility are unbiased. In contrast, the results for BS ATM implied volatility and the EGARCH model are mixed, such that no clear conclusion regarding their forecasting bias is possible. Further, the findings for the encompassing regressions indicate that the volatility forecasts based on BS ATM implied volatility and the GARCH models are unbiased. Despite these findings, Day and Lewis (1992) report that their out-of-sample results do not allow them to draw additional conclusions regarding the relative predictive power of implied volatility relative to GARCH models.[12] The controversy surrounding the results of Canina and Figlewski (1993) and Christensen and Prabhala (1998) is discussed below.

Similar to Day and Lewis (1992), Canina and Figlewski (1993) also examine the information content of implied volatility using S&P 100 index options.[13] However, while the analysis of Day and Lewis (1992) is based on BS ATM implied volatilities from short-term S&P 100 index options, Canina and Figlewski (1993) analyse the forecasting performance of implied volatility for different maturities and intrinsic value groups. They argue that taking the average implied volatility from options with different intrinsic values and maturities is misleading, because the existence of the volatility smile and the volatility term structure indicate systematic differences across implied volatilities. By estimating the univariate regression equation (5.1) for each maturity/intrinsic value group separately, they find that all intercepts are significantly different from zero and that the estimated coefficients on implied volatility do not differ significantly from zero in most subsamples.[14] Further, they estimate encompassing regressions for each group and document that the coefficients on historical volatility are generally significantly positive. In contrast, the coefficients on

[12]See Day and Lewis (1992), pp. 281-286.

[13]Their sample comprises the period from March 1983 to March 1987.

[14]To capture serial correlation due to overlapping samples, they estimate the above regressions using GMM estimator. Yu et al. (2010) note that serial correlation in overlapping samples yields an underestimated standard error for the coefficient on historical volatility. Thus, empirical studies that ignore serial correlation from overlapping samples are biased towards confirming the result that "historical volatility provides an efficient forecast of future volatility". See Yu et al. (2010), p. 3.

implied volatility are typically not significantly different from zero. Overall, Canina and Figlewski (1993) conclude that implied volatility has no explanatory power for future volatility and does not include the information contained in historical volatility. In contrast, their results provide evidence that historical volatility contains information about future volatility. As an explanation for their results, they suggest that implied volatility is affected by additional factors that are not included in traditional option pricing theory (e.g., market frictions, liquidity constraints, investor preference for certain option payoffs etc.).[15]

Because the findings of Canina and Figlewski (1993) do not agree with those of most previous studies (e.g., Day and Lewis (1992) and Lamoureux and Lastrapes (1993)), subsequent studies have critically analysed Canina and Figlewski's (1993) results. For instance, in an oft-cited article, Christensen and Prabhala (1998) re-examine the results obtained Canina and Figlewski (1993) by estimating regression (5.3) using instrumental variables. By applying instrumental variables to estimate the regression they correct for the errors-in-variables problem in implied volatility that can be induced by dividend payments, non-synchronous data, bid-ask spreads, and misspecification of the option pricing model employed. After the correction, they report that BS ATM implied volatility exhibits lower prediction bias than that documented by Canina and Figlewski (1993). In addition to addressing the errors-in-variables problem, Christensen and Prabhala (1998) contend that the downward bias can be explained by the longer time series that they consider, which includes a regime shift near the October 1987 crash.[16] In particular, they observe a reduced forecasting bias for implied volatility following the crash.[17,18] Further, they attribute the divergent results obtained in their study and Canina and Figlewski (1993) to different sampling frequencies. Specifically, they adopt a lower (monthly) sampling frequency than Canina and Figlewski (1993), which allows them to construct non-

[15] See Canina and Figlewski (1993), pp. 670-768.

[16] Their sample contains monthly S&P 100 index options data from November 1983 to May 1995.

[17] See Christensen and Prabhala (1998), p. 127.

[18] Ederington and Guan (2002) also argue that the discrepancies between Christensen and Prabhala (1998) and Canina and Figlewski (1993) are due to the former's inclusion of the October crash 1987 while the latter include it. See Ederington and Guan (2002), p. 37.

overlapping samples.[19] Overall, they find that BS ATM implied volatility provides better volatility forecasts than historical volatility and, in some cases, subsumes all information contained in historical volatility.[20]

Similar to Canina and Figlewski (1993), Ederington and Guan (2005) also examines the information content of implied volatility for options with different strike prices. Their sample contains daily settlement prices for short-term S&P 500 futures options from January 1988 to March 2003. Before performing the analysis for different strike groups, they estimate regression models (5.1) and (5.3) based on the complete (pooled) data set.[21] They report that the estimated intercepts and coefficients of implied and historical volatility are significantly different from zero at the 1% level. Thus, this finding does not confirm the hypothesis that implied volatility is an unbiased and efficient estimator of future volatility.[22] To investigate the predictive ability of implied volatility across different strike prices and option types (put and call options), they separately estimate the regressions presented above for different moneyness levels and option types. The estimation results of the univariate regressions (5.1) indicate that implied volatility from options with moderately high strike prices is an unbiased predictor of future volatility. Further, the estimated coefficients of the implied volatility variables for these options are not significantly different from one.[23] In contrast, they reject the hypothesis that volatility forecasts based on implied volatility are unbiased for options with low strikes, ATM strikes, and high strikes. Thus, they conclude that the information content of implied volatility is related to the strike price, whereby moderately OTM calls and ITM puts provide the most information concerning future volatility. Their findings from the encompassing regressions provide evidence that for most moneyness classes, implied volatility

[19] According to Christensen and Prabhala (1998) overlapping samples produce autocorrelated errors that lead to imprecise and inconsistent regression estimates. See Christensen and Prabhala (1998), p. 129.

[20] See ibid., p. 148.

[21] They use the sample standard deviation or, alternatively, the GARCH(1,1) model to forecast volatility based on past returns.

[22] See Ederington and Guan (2005), p. 1438.

[23] Moreover, they obtain the highest relative explanatory power for options with moderately high moneyness levels.

subsumes the information contained in historical volatility. Thus, the above results based on the pooled data set change substantially when the regression analysis is performed separately for each moneyness/option type group.[24]

In a related study, Fleming et al. (1995) investigate the statistical properties and predictive power of S&P 100 implied volatility for stock market volatility.[25] Because their analysis of the statistical properties suggests that implied and historical volatility follow near random walk processes, they subtract lagged implied volatility from both sides of regression equation (5.1) to avoid the spurious regression problem.[26] Based on the adjusted regression, they report that the implied volatility forecasts are biased. They suggest that the forecasting bias is due to the misspecification of the option pricing model, neglecting the wildcard option embedded in the S&P 100 index option contract, and infrequent trading. Further, they perform additional orthogonality tests and find that historical volatility does not provide additional information relative to S&P 100 implied volatility for the prediction of volatility.[27] In summary, their results support the prior findings of Christensen and Prabhala (1998).

Most of the above-cited studies indicate that implied volatility is a biased predictor of future volatility. This finding suggests that implied volatility may not contain all information on future volatility or that the information is not processed correctly by the option pricing model considered. Because determining the source of the error is essential for generating improved volatility forecasts based on implied volatility, the following Section presents selected studies that investigate whether this forecasting bias can be attributed to measurement errors in implied volatility.

[24] See Ederington and Guan (2005), pp. 1442-1450.
[25] They consider the sample period from January 1986 to December 1992.
[26] Their methodology follows Fleming (1998). In addition to adjusting the regression equation, they use the GMM estimator to account for potential residual heteroskedasticity and autocorrelation.
[27] See Fleming et al. (1995), pp. 290-300.

5.2.2. The Errors-in-Variables Problem Due to Measurement Errors in Implied Volatility

In general, measurement errors in the independent variables affect OLS coefficient estimates. Implied volatility can differ from the market's true volatility expectation for several reasons. This Section considers the effects of measurement errors in implied volatility that are induced by market imperfections (e.g., non-synchronous data, and bid-ask spreads). Other error sources, e.g., misspecification of the option pricing model, are discussed in Chapter 5.2.4.

Christensen and Prabhala (1998) as well as Figlewski (1997) describe the potential effects of measurement errors on the forecasting performance of implied volatility. According to Christensen and Prabhala (1998), the implied volatility derived from an option pricing model contains the true implied volatility plus measurement error.[28] Thus, using implied volatility to forecast volatility without accounting for measurement errors is misleading, because this volatility forecast does not reflect the correct market information concerning future volatility.[29] Figlewski (1997) provides a further explanation. According to Figlewski (1997), arbitrage trades between stock index options and the underlying stock index are expensive and risky. Therefore, arbitrage trading that exploits the difference between current market implied volatility and traders' expectations of future volatility becomes increasingly difficult. As a consequence, some of these expectations are not incorporated into option prices and the hypothesis that implied volatility contains information on future volatility is rejected.[30] Ultimately, both explanations lead to the same result: implied volatility does not contain the correct market information concerning future volatility.

Technically, the errors-in-variables (EIV) problem causes the slope coefficient of implied volatility to be downward biased in the above regressions (see equations (5.1) and (5.3)). Further, if implied and ex-post volatility are positively correlated

[28] This hypothesis requires that the applied option pricing model holds.
[29] See Christensen and Prabhala (1998), pp. 136-137.
[30] See Figlewski (1997), p. 64.

and the slope parameter of implied volatility is positive, the slope coefficient of ex-post volatility is upward biased. Thus, the OLS estimates of regression equations (5.1) and (5.3) are inconsistent and the conclusion that implied volatility is biased and inefficient is misleading.[31]

Christensen and Prabhala (1998) contend that the OLS estimation results in their study from regressions (5.1) and (5.3) are subject to the EIV problem, as the implied volatility of S&P 100 index options is affected by measurement errors. To obtain consistent estimates in the presence of the EIV problem, they employ the instrumental variables method. Applying instrumental variables, they provide evidence that forecasts based on implied volatility are unbiased and efficient. Drawing from these results, they conclude that the forecast bias of implied volatility obtained via OLS estimation can be attributed to the EIV problem. However, as several sources of measurement error (e.g., early exercise, dividends, non-synchronous data, bid-ask spread, the wild-card option, and the misspecification of the BS model) can generate the EIV problem in S&P 100 implied volatility, Christensen and Prabhala (1998) fail to identify the key source(s) of error.[32]

In addition, Ederington and Guan (2002) note that Canina and Figlewski (1993) consider the implied volatility for each strike price separately and do not average out measurement errors across different strike prices as is done in most other studies. Further, they point out that S&P 100 index options are affected by the non-synchronous data problem[33] and that existing transaction costs hamper arbitrage trading between markets. They suggest that this can explain the relatively low coefficient values of implied volatility in the encompassing regressions documented by Canina and Figlewski (1993).[34] To reduce the effects of measurement errors, Eder-

[31]See Christensen and Prabhala (1998), p. 137.
[32]See ibid., pp. 136-140.
[33]This problem is caused by the different closing times of the options market and the New York Stock Exchange.
[34]See Ederington and Guan (2002), pp. 33-34.

ington and Guan (2002) consider S&P 500 futures options, which are less affected by measurement errors than S&P 100 index options.[35]

Similar to Christensen and Prabhala (1998), they apply the instrumental variables method to estimate the encompassing regressions. They observe smaller differences between the OLS and instrumental variable coefficients than those documented by Christensen and Prabhala (1998).[36] Their results for the full sample provide evidence that implied volatility is an unbiased and efficient predictor of future volatility.[37] Because Ederington and Guan (2002) report smaller differences between the OLS and instrumental variable coefficients for S&P 500 futures options, their findings support the above results of Christensen and Prabhala (1998).[38] This discussion regarding the impact of measurement errors in implied volatility demonstrates that a comparison of different volatility forecasting models should be based on synchronous data that are taken from liquid markets with low transaction costs. Otherwise, the forecast evaluation results are influenced by the effects outlined above.

Having analysed the impact of measurement errors on the regression results, the next Section considers the effects of using intraday returns to estimate daily spot volatility.

5.2.3. Effects of Using Intraday Returns as an Ex-Post Volatility Measure

The increasing availability of high-frequency data stimulated the development of new volatility proxies, called realised volatility measures, which are based on intraday data.[39] The idea of using high-frequency data dates back to Merton (1980), who

[35] S&P 500 futures options and corresponding futures are traded in tandem, such that synchronous quotes are available and arbitrage is straightforward.

[36] Their analysis is based on S&P 500 ATM futures options from January 1983 to September 1995.

[37] However, when the 1987 crash is excluded from the data set, the hypothesis that implied volatility is unbiased and efficient is rejected.

[38] See Ederington and Guan (2002), pp. 41-45.

[39] Following the underlying literature, the expression "realised volatility" is used in the following to denote daily return volatility that is based on high-frequency returns to avoid confusion in comparisons with the literature.

demonstrates that if the variance is estimated by the sum of squared returns over a fixed period and the sampling frequency can be increased arbitrarily, the estimator converges towards the true volatility.[40] However, market microstructure noise can induce severe bias in the estimated daily volatility when the sampling frequency is too high.[41]

Andersen and Bollerslev (1998) examine the effect of ex-post volatility measures on the evaluation of forecasts from standard volatility models. They exhibit that traditional forecast evaluation criteria indicate poor forecasting performance for standard volatility models when volatility is measured based on daily squared returns.[42] This poor predictive power can be explained by the use of daily squared returns as an ex-post volatility measure, which provides a noisy estimate of latent volatility. For this reason, Andersen and Bollerslev (1998) suggest an alternative measure of ex-post volatility based on intraday returns.[43]

While Andersen and Bollerslev (1998) focus on evaluating of the general predictive power of standard volatility models, Poteshman (2000) uses realised volatility to compare the forecasting performance of implied and historical volatility for S&P 500 index options.[44] In particular, he analyses the influence of three different ex-post volatility measures on the forecast bias and informational efficiency of implied volatility. Two of the three volatility measures are calculated based on daily squared returns using daily closing prices, respectively the daily 3:00 PM index level.[45] The third volatility measure is defined as the daily sum of squared five-minute S&P 500 returns. By estimating regression equation (5.1) for each ex-post volatility measure, he reports that the estimated intercepts decrease towards zero and the slope coef-

[40] Additionally, he assumes that the market returns follow a diffusion-type stochastic process, the mean and variance of which are constant or at least change slowly over time. See Merton (1980), pp. 355-359.
[41] See McAleer and Medeiros (2008), p. 12.
[42] They report low R^2 values, despite that the volatility models are correctly specified.
[43] See Andersen and Bollerslev (1998), p. 886.
[44] Here, historical volatility is calculated from five-minute S&P 500 returns for different fixed time periods (e.g., one month) up to the present. The implied volatility is computed using the BS model from near ATM call options.
[45] The index level is derived from S&P 500 futures transaction data.

ficients of implied volatility increase towards one when moving from daily squared returns to realised volatility.[46] Similar to the results of Andersen and Bollerslev (1998), he finds that the predictive power of implied volatility is higher when using realised volatility as a volatility measure.[47] Furthermore, he demonstrates that in the full sample from June 1988 to August 1997, nearly half of the forecasting bias of S&P 500 implied volatility disappears if ex-post volatility is measured based on squared five-minute returns.[48]

Similar to Poteshman (2000), Blair et al. (2001) also compare the information content of implied and historical volatility for forecasting volatility using different ex-post volatility measures.[49] In particular, they calculate ex-post volatility based on squared excess returns and the sum of squared five-minute returns. The implied volatilities are obtained from implied volatility index.[50] In addition to the VIX, they investigate the forecast information contained in the volatility forecasts obtained from the GJR-GARCH model, the lagged sample standard deviation, and lagged realised volatility. They find that the VIX is more informative than the individual forecasts based on historical volatility.[51] This result does not depend on the forecast horizon and the ex-post volatility measure considered.[52,53] Moreover, they conclude that historical volatility provides little additional information beyond the VIX for one-day-ahead forecasts and that the VIX contains all relevant information for longer forecasting periods.[54]

[46] While the intercepts are not significantly different from zero regardless of the volatility measure, the slope coefficient of implied volatility is only significantly different from one when measuring ex-post volatility based on daily closing prices.

[47] These results refer to the sample period from June 1993 to August 1997.

[48] See Poteshman (2000), pp. 18-29.

[49] Their sample comprises the period from January 1987 to December 1999.

[50] In particular, they use VIX data computed by the CBOE. The implied volatilities entering the VIX are calculated from a binomial model.

[51] Their comparison is based on the ranked R^2 value for equation (5.1) or (5.2).

[52] They evaluate the performance of one-, five-, ten-, or twenty-days-ahead volatility predictions.

[53] As in Andersen and Bollerslev (1998) they report higher R^2 values for the regressions when realised volatility is applied to measure ex-post volatility instead of using daily squared returns.

[54] This finding is based on the comparison of R^2 values between univariate and multivariate regression models. Note that the coefficient of determination always increases if an additional regressor with an associated t-statistic greater (less) than one (negative one) is added to the regression equation. Thus, this conclusion for one-day-ahead forecasts reported by Blair et al. (2001) is questionable.

The following Section examines the effect of the option pricing model used to compute implied volatility on forecasting performance.

5.2.4. Implications of the Choice of Option Pricing Model

In addition to investigating the effect of alternative ex-post volatility measures on the evaluation of volatility forecasts, Poteshman (2000) examines whether the forecasting bias of implied volatility documented by several studies[55] is due to the misspecification of the applied option pricing model. In particular, he compares the forecasting results from using the standard BS model and the Heston (1993) model, which allows for a non-zero price of volatility risk and non-zero correlation between innovations and the level as well as volatility of the underlying asset. The estimation results for equation (5.3) indicate that the volatility forecasts for the S&P 500 index based on Heston implied volatility are unbiased and efficient.[56] In addition, he conducts a simulation experiment that supports this finding.[57] Overall, Poteshman (2000) demonstrates that the forecasting bias of implied volatility diminishes when ex-post volatility is measured based on five-minute returns and an option pricing model that allows for a volatility risk premium is used to derive implied volatility.

Similar to Poteshman (2000), Shu and Zhang (2003) also employ the BS and Heston (1993) models to investigate the effect of model misspecification on the forecasting ability of implied and historical volatility.[58] In their analysis, they calibrate both option pricing models to daily S&P 500 index option prices from January 1995 to December 1999. Estimating univariate regressions of type (5.1), they find that Heston implied volatility has lower explanatory power and greater forecasting bias than

[55]See, for example, Fleming et al. (1995), among others.

[56]This finding is obtained by using five-minute futures data to measure ex-post volatility. Because similar results are presented for the BS model, this finding alone is not sufficient to draw the conclusion that the forecasting bias of BS implied volatility is due to misspecification errors.

[57]See Poteshman (2000), p. 18.

[58]As in Poteshman (2000), they also use different ex-post volatility measures to examine whether measurement errors in ex-post volatility affect the forecasting performance of implied and historical volatility.

BS implied volatility.[59] Further, encompassing regressions provide evidence that volatility forecasts based on Heston and BS implied volatility subsume all information contained in historical returns. The multivariate regressions also demonstrate that the implied volatility calculated from the Heston (1993) model has less explanatory power in forecasting S&P 500 index volatility than BS implied volatility. Shu and Zhang (2003) suppose that the inferior forecasting performance of the Heston (1993) model is due to a mismatch between the model's underlying stochastic process and the true data-generating process. While Shu and Zhang (2003) concur with Poteshman (2000) that implied volatility outperforms historical volatility when forecasting S&P 500 index volatility, Shu and Zhang's (2003) findings with respect to the forecasting bias of implied volatility contradict those of Poteshman (2000). Although Shu and Zhang (2003) cite Poteshman (2000) and both articles use S&P 500 index options data, the former do not provide an explanation for the divergent findings. Comparing the structures of the two studies suggests that the contradicting results are may be due to different sample periods, different calibration methods for the Heston model, and/or different underlyings assets used to derive the implied volatilities from S&P 500 index options.[60]

Whereas Poteshman (2000) and Shu and Zhang (2003) present empirical results regarding whether volatility forecasts based on implied volatility are biased, the analysis of Chernov (2007) provides theoretical arguments regarding this "(un)biasedness puzzle". By using affine jump-diffusion models that allow for stochastic volatility with independent jumps in returns and volatility, he shows that the forecasting bias of implied volatility can be explained by volatility and jump risk.[61] In particular, he demonstrates that the decomposition of the volatility forecast into BS implied volatility and the volatility risk premium implies the standard encompassing regression presented above (see Section 5.1) that includes an additional term to capture volatility risk premia. Chernov (2007) demonstrates that this additional term is

[59] This finding holds for various ex-post volatility measures. See Shu and Zhang (2003), pp. 88-89.

[60] While Poteshman (2000) considers SPX futures data, Shu and Zhang (2003) base their study on daily S&P 500 index closing prices.

[61] He assumes that these additional risk factors are not diversifiable, such that investors demand related risk premia.

a linear function of spot volatility and suggests estimating it by using the high-low range-based estimator. If this additional term is not included in the regression equation, then the slope coefficient of implied volatility is downward biased, as spot volatility and implied volatility are correlated.[62] Further, because the estimation of the additional regressor introduces an errors-in-variables problem, he recommends a GMM framework to address this.[63]

To test his theoretical results, he estimates the standard regressions described above with and without the inclusion of spot volatility, which is estimated by the high-low range-based estimator as an additional predictive variable. He analyses options data on the S&P 100 index, the National Association of Securities Dealers Automated Quations (NASDAQ) 100 index, and three foreign exchange rate series.[64] The OLS estimation results of the regressions indicate that implied volatility is generally a biased predictor of realised volatility. This finding agrees with the majority of the literature. In contrast, the GMM results provide evidence that the implied volatility bias disappears after volatility risk premia are accounted for in the regression equation.[65] Overall, Chernov (2007) provides theoretical and empirical arguments that support the above mentioned results of Poteshman (2000).

Because the above studies using ATM options do not consider the information provided by OTM and ITM options, Jiang and Tian (2005) suggest the use of model-free implied volatility to capture this additional information. This approach is based on cross-sectional options prices and is described in detail in Chapter 4.[66] They compare the forecasting performance and information content of model-free implied volatility

[62]See Chernov (2007), pp. 412-414.

[63]See ibid., p. 420.

[64]The S&P 100 (NASDAQ 100) index data refer to the period from January 1986 (January 1995) to June 2001 (June 2001). The three foreign exchange rate series cover the period from October 1984 to June 2001.

[65]See Chernov (2007), pp. 417-420.

[66]In addition, they argue that model-free implied volatility does not depend on a particular option pricing model. This implies the following advantage: while tests based on BS implied volatility are joint tests of market efficiency and the applied option pricing model, model-free implied volatility represents a direct test of market efficiency. See Jiang and Tian (2005), pp. 1305-1308.

with BS ATM implied volatility and historical volatility. Their sample contains S&P 500 index options data from June 1988 to December 1994.[67]

In accordance with previous studies, using univariate regressions, they demonstrate that BS implied volatility explains more of the variation in future realised volatility than lagged realised volatility. However, they find no evidence that BS implied volatility is an efficient forecast of future volatility. In contrast, when using model-free volatility to forecast volatility, they demonstrate that model-free volatility subsumes all information provided by BS ATM implied volatility and historical volatility.[68] Thus, they argue that model-free volatility represents a more efficient forecast of future volatility. They conclude that their results support the informational efficiency of the options market. Further, they find that model-free volatility provides better forecasting results than BS ATM and historical volatility.[69] Their results are robust to alternative estimation methods, different samples, and different measures of realised volatility.[70]

While most empirical studies examine volatility forecasts extracted from stock index options, Taylor et al. (2010) analyse the information on volatility contained in the stock options of individual firms. They use daily closing option quotes and stock prices for 149 US firms from January 1996 to December 1999 to compare volatility forecasts based on BS ATM implied volatility, model-free volatility, and historical volatility.[71] Similar to studies cited above, they consider option-life forecasts (non-overlapping monthly periods) and estimate univariate regressions for each volatility forecast and individual firm.[72]

They report that implied volatility forecasts provide more information on future volatility than historical volatility for more than 85% of the firms considered. With

[67] Specifically, they use tick-by-tick data to reduce measurement errors. See Jiang and Tian (2005), pp. 1318-1319.
[68] They use encompassing regressions to examine the information content of model-free volatility.
[69] See ibid., pp. 1323-1329.
[70] See ibid., pp. 1329-1336.
[71] In particular, they estimate a GJR(1,1)-MA(1,1) model to capture the asymmetric volatility effects of historical stock returns.
[72] See Taylor et al. (2010), pp. 871-873.

respect to the predictive power of BS ATM implied volatility and model-free volatility, they find that volatility forecasts based on BS ATM implied volatility are more informative than using model-free volatility for more than 50% of the firms in their sample. Further, the estimation results of the encompassing regressions indicate that historical volatility typically contains additional information not conveyed by implied volatility. Additionally, neither implied volatility variable subsumes all of the information covered by the other variable. A cross-sectional analysis indicates that the forecasting performance of implied volatility forecasts that is observed for individual firms depends on firm-specific option liquidity. Greater predictive power is observed for firms with more liquid options.[73]

The findings of Taylor et al. (2010) contrast with the above mentioned results of Jiang and Tian (2005). Thus, although model-free volatility is based on a larger information set than BS ATM implied volatility, the higher measurement errors of OTM options and relatively low trading volumes of individual stock options mean that model-free volatility exhibits worse performance. Taylor et al. (2010) conclude that although model-free volatility represents a theoretically appealing approach, the cross-sectional information content of individual option prices is outweighed by the illiquidity of OTM options.

Having presented empirical results concerning the predictive ability of alternative option pricing models, the next Section considers whether time series models that account for long memory effects provide volatility forecasts superior to those obtained using implied volatility.

5.2.5. Volatility Forecasts from Long Memory Models

Most of the previously cited studies compare the forecasting power of implied volatility with volatility forecasts based on lagged standard deviations or GARCH models. As intraday data have become increasingly available, new volatility measures have

[73]See ibid., pp. 875-880.

been developed and suitable time series models have been suggested to produce volatility forecasts.[74] The studies of Li (2002), Martens and Zein (2004), and Becker et al. (2006) compare the forecasting performance of implied volatility and ARFIMA models (or combinations thereof) and are described in the following.

Li (2002) presents the first study that provides an analysis of the forecasting ability of long memory models using realised volatility and implied volatility. He examines high-frequency returns on the German Deutsche Mark, the Japanese yen, the British pound, and the US dollar from December 1986 to December 1998.[75] Further, he considers daily BS ATM implied volatilities for OTC forward currency options with fixed maturities for different sample periods that are directly observed in the market. Similar to the findings of Andersen et al. (2001b) for the realised volatilities of the 30 stocks in the Dow Jones Industrial Average, Li (2002) documents long memory effects in the realised volatility series and fits an ARFIMA process to each series. He reports that volatility forecasts based on long memory models provide additional information relative to option-implied volatility forecasts at horizons of from one month to six months. In addition, long memory models provide better volatility predictions than option-implied volatility at longer forecast horizons. In summary, he concludes that volatility forecasts can be improved by combining forecasts based on long memory models and option-implied volatility.[76]

While Li (2002) studies different exchange rate series, Martens and Zein (2004) compare the forecasting performance of long memory models and option-implied volatility for different asset classes.[77] Specifically, they provide results for the S&P

[74] As mentioned above, Andersen and Bollerslev (1998) introduced the realised volatility measure, which is based on the sum of squared intraday returns. Subsequently, Andersen et al. (2003) suggested the ARFIMA process to model the long memory of realised volatility.

[75] He constructs realised volatility series based on the daily sum of squared five-minute returns.

[76] See Li (2002), pp. 9-25.

[77] Another notable difference between the studies of Martens and Zein (2004) and Li (2002) concerns the overlapping data problem. Whereas Li (2002) corrects for downward-biased OLS standard errors in the forecast evaluation regressions that are induced by overlapping data using the Hansen (1982) variance-covariance matrix, Martens and Zein (2004) use non-overlapping data.

500 index, YEN/USE exchange rate, and Light, Sweet Crude Oil.[78] Their sample comprises transaction data for S&P 500 futures and daily data for S&P 500 index futures options from the beginning of 1996 to the end of 2000.[79] They also compute realised volatility based on the sum of squared five-minute returns.[80]

With respect to the S&P 500 volatility forecasts, they report that implied volatilities outperform the GARCH(1,1) model, which agrees with the findings of Christensen and Prabhala (1998) and Fleming (1998). However, the use of time series models that account for long memory effects affects the relative performance results of time series models and implied volatilities. Martens and Zein (2004) demonstrate that volatility forecasts based on long memory models provide similar and in some cases better prediction results than implied volatility. Furthermore, they report that S&P 500 volatility forecasts based on long memory models and implied volatility contain additional information on future volatility beyond the information covered by the other approach. Thus, according to Martens and Zein (2004), a combined forecast approach using implied volatility and the ARFIMA forecast should improve predictions of S&P 500 volatility.[81] In this respect, the study of Martens and Zein (2004) confirms the results of Li (2002) for the S&P 500.

Similar to Martens and Zein (2004) and Li (2002), Becker et al. (2006) investigate whether implied volatility reflects all information provided by alternative model-based volatility forecasts. They extend the analysis of the previous studies by using a wider set of conditioning information.[82] To adapt the findings of Jiang and Tian (2005), they apply a publicly available implied volatility index, the VIX, which is

[78] Because the existing thesis investigates DAX volatility forecasts, the results for the S&P 500 are presented in the following. See Martens and Zein (2004) for further results regarding the other asset classes.

[79] Martens and Zein (2004) apply the quadratic approximation method for American options developed by Barone-Adesi and Whaley (1987) that uses the generalised BS formula.

[80] See Martens and Zein (2004), pp. 1005-1008.

[81] See ibid., pp. 1019-1027.

[82] While the studies above compare the forecasting power of implied volatility using either GARCH or ARFIMA models, Becker et al. (2006) apply a range of alternative volatility forecasts including GARCH, ARFIMA, and stochastic volatility models.

based on the concept of Britten-Jones and Neuberger (2000).[83] Further, Becker et al. (2006) use high-frequency returns to measure realised volatility and follow Chernov's (2007) suggestion concerning the consideration of volatility risk.[84] Their testing strategy is based on two approaches: first, they apply the framework suggested by Fleming (1998) that allows the researcher to examine the orthogonality of implied volatility forecast errors with respect to a particular information set. Second, they implement the forecast-encompassing tests developed by Harvey and Newbold (2000) that can be used to test whether implied volatility encompasses one or more alternative volatility forecasts.[85] The set of alternative prediction models contains, among others, the GARCH(1,1) model, an asymmetric GJR-GARCH model, a stochastic volatility model, ARMA and ARFIMA models based on realised volatility, and exponentially weighted moving averages of squared returns.

Their sample contains daily VIX data and high-frequency S&P 500 index returns from January 1990 to October 2003. In general, the forecast-encompassing tests provide evidence that the alternative prediction models provide information that is correlated with the VIX forecast errors. While the orthogonality tests for the daily sampling scheme generally reject the hypothesis that implied volatility subsumes all information, the test results when using monthly sampling confirm the hypothesis. Becker et al. (2006) suggest that these contradictory findings with respect to sampling frequency can be explained by size distortions in the tests and different sample sizes.[86] Becker et al. (2006) summarise that volatility forecasts based on the VIX

[83] To provide a more practical standard for trading and hedging, the CBOE revised the methodology of the VIX in 2003. Thereafter, the VIX is calculated based on the concept of Britten-Jones and Neuberger (2000) by using options on the S&P 500 index with strike prices near the current index level and maturities close to 22 trading days. See Becker et al. (2006), p. 140.

[84] To capture the impact of the volatility risk premium, Chernov (2007) suggests the inclusion of an additional term in the encompassing regression used to evaluate the volatility forecasts. See the previous Section.

[85] These tests allow the researcher to examine whether the forecast error of implied volatility can be explained by the forecast errors of alternative prediction models. To answer this research question, the forecast errors of implied volatility are regressed on the forecast errors of alternative models (see Becker et al. (2006) for a more detailed description of the test concept). Becker et al. (2006) implement the orthogonality tests as a supplement to the encompassing regressions, because these tests support a more general analysis.

[86] In principle, overlapping data can induce size distortions. However, Becker et al. (2006) employ specific techniques to address the overlapping data problem.

can be improved by using additional information, e.g., volatility forecasts from time series models.[87]

Overall, the studies of Li (2002), Martens and Zein (2004), and Becker et al. (2006) suggest combining implied volatility and model-based forecasts, e.g., long memory models.

5.2.6. Empirical Studies Evaluating Volatility Forecasting Performance Based on Loss Functions

The review of the literature provided above considers the evaluation of volatility forecasts based on encompassing regressions. This Section provides an overview of studies using loss functions to assess the performance of different volatility forecasting methods.

Fung and Hsieh (1991) present an early paper comparing the forecasting performance of implied volatility and standard volatility forecasts based on historical prices and use high-frequency data to measure realised volatility. In their study, they evaluate S&P index volatility forecasts using two loss functions, the root mean square error and the mean absolute error, where realised volatility is calculated as the daily standard deviation of 15-minute returns. Their samples contain tick-by-tick data for S&P futures and options contracts for the period from March 1983 to July 1989.[88] The one-day-ahead volatility forecasts are generated using the rolling standard deviation based on daily closing prices, the standard deviation based on the extreme value method suggested by Parkinson (1980), and two implied volatility series for S&P ATM calls and puts.[89] In addition, they construct two volatility forecast series by defining a random walk model and fitting an autoregressive model to the realised volatility series. They find that the prediction errors of the volatility forecasts based on implied volatility are similar to the errors of the volatility forecasts based on the

[87] See Becker et al. (2006), pp. 145-152.
[88] In addition to analysing the stock market, they also consider the bond and currency markets.
[89] The implied volatilities are computed using the Barone-Adesi and Whaley (1987) approximation.

historical volatility series. Thus, they conclude that implied volatility from options on S&P futures does not contain additional information on future S&P volatility relative to the standard volatility forecasts based on historical prices.[90]

While Fung and Hsieh (1991) compare option-implied volatility forecasts and simple volatility forecasts based on historical prices, Gospodinov et al. (2006) investigate an extended set of alternative volatility forecasting models. By performing an analysis of the time series properties of S&P 100 volatility, they find evidence of slow mean-reverting behaviour and long memory effects.[91] To reproduce these characteristics, they fit an ARFIMA model, a near-integrated autoregressive model developed by Gospodinov (2002), and a Fractionally Integrated Exponential GARCH (FIEGARCH) model to the data.[92] The predictive ability of the different volatility forecasting methods is examined across different loss functions where daily returns on the S&P 100 index are used to proxy for the latent volatility process.[93] Their evaluation results for one-step-ahead volatility forecasts reveal that implied volatility contains useful information on future volatility. Moreover, they report that forecast combinations based on alternative volatility models tend to provide lower prediction errors than using the individual models. However, their forecast evaluation approach does not provide information concerning whether the prediction errors from the combined forecasts are significantly lower than those from the individual models.[94]

By using the SPA test developed by Hansen (2001), Koopman et al. (2005) investigate the relative forecasting performance of various volatility models for the S&P 100 index. The SPA test permits a formal examination of whether a benchmark model is significantly outperformed by a set of alternative models. The relative forecasting performance used to compare the volatility models is measured by tak-

[90] See Fung and Hsieh (1991), pp. 1-18.
[91] Their data set contains daily data for the VIX and the S&P 100 index from June 1988 to May 2002.
[92] In addition, they estimate an EGARCH model, a stochastic volatility model, and combined forecasts.
[93] See Gospodinov et al. (2006), pp. 381-387.
[94] See ibid., pp. 393-397.

ing the differences of the forecast error series for each model combination. Similar to Li (2002), Martens and Zein (2004), and Becker et al. (2006), Koopman et al. (2005) also estimate daily or realised volatility as the sum of squared five-minute returns on the S&P 100. The forecast errors are calculated for different loss functions (mean square error (MSE), mean absolute error (MAE), mean square error adjusted for heteroskedasticity (HMSE), and mean absolute error adjusted for heteroskedasticity (HMAE)) based on this realised volatility measure. They compare the forecasting performance of unobserved Autoregressive Moving Average (ARMA) components models, ARFIMA models, stochastic volatility models, and GARCH models. To account for recent research findings, they add realised and implied volatilities as explanatory variables to the stochastic volatility and GARCH models.[95] The empirical investigation is performed for one-step-ahead forecasts for the period from January 1997 to November 2003.[96]

Comparing the forecasting performance of the volatility models based on the loss functions, they find that the volatility forecasts provided by the unobserved ARMA components and the ARFIMA models outperform the predictions from the stochastic volatility and GARCH models. They identify two reasons for the relatively poor prediction results of the GARCH and stochastic volatility models: first, they refer to volatility measurement errors in daily squared returns, and second, they argue that stochastic volatility and GARCH models react slowly to volatility changes. Within the class of volatility models using daily returns, they report that the stochastic volatility model extended by either lagged realised or implied volatility provides the best forecasting results. Among the volatility models based on realised volatility, the ARFIMA model produces the most accurate volatility forecasts. These results are supported by the evidence provided by the SPA test.[97]

Similar to Koopman et al. (2005), Martin et al. (2009) also compare the predictive ability of options-based and return-based volatility forecasts by performing the SPA

[95]The option-implied volatility is measured by the VIX.
[96]See Koopman et al. (2005), pp. 445-448.
[97]See ibid., pp. 465-472.

test. However, the latter paper places greater emphasis on the influence of different noise-corrected volatility measures on the forecast evaluation results. While most empirical studies employ the sum of squared returns over small, regular intervals to proxy for unobserved volatility, they calculate a range of realised volatility measures that differently account for empirical regularities of microstructure noise.[98] Another focus of their study is the robustness of the relative performance of options-based versus return-based volatility forecasts with respect to the choice of the option pricing model. Therefore, they use ATM implied volatility and model-free implied volatility to derive options-based volatility forecasts.[99] The performance of the options-based forecasts is compared to selected return-based volatility forecasts for three DJIA stocks and the S&P 500 index. To capture the empirical features of the stocks and the index, they fit an ARFIMA model, various GARCH models, and, for completeness, an ARMA model to the data.[100] By using the SPA test, where option-implied volatility represents the benchmark model (respectively, ATM implied volatility or model-free implied volatility), they investigate whether options-based forecasts are outperformed by a set of alternative models.[101]

With respect to a one-day-ahead and a 22-day-ahead forecast horizon, they document that the model-free implied volatility exhibits poor prediction results for the three individual stocks and the S&P 500 index.[102] In contrast, the SPA test reveals that ATM implied volatility provides superior volatility forecasts for the three DJIA stocks.[103] However, volatility forecasts based on ATM implied volatility do not outperform return-based models with respect to the S&P 500 index. These find-

[98] Specifically, they use the two-scale realised volatility estimator suggested by Zhang et al. (2005) and Aït-Sahalia et al. (2011), the realised kernel estimator developed by Barndorff-Nielsen et al. (2008), the optimal sampled realised volatility estimator proposed by Bandi and Russell (2006), the bi-power variation measure of Barndorff-Nielsen and Shephard (2004), the modified alternation estimator developed by Barndorff-Nielsen and Shephard (2007), and the standard realised volatility measure based on squared five-minute returns.

[99] The implied volatilities for the European-style index options are computed via the BS option pricing model and for the three Dow Jones Industrial Average (DJIA) stocks using a binomial tree method.

[100] Their sample contains intraday spot and options prices from June 1996 to June 2006.

[101] See Martin et al. (2009), pp. 77-79.

[102] The SPA test indicates that at least one model in the set of alternative models significantly outperforms model-free implied volatility.

[103] This particularly holds for the 22-day-ahead forecasts.

ings hold regardless of the realised volatility measure considered and, thus, do not depend on the method applied to correct for microstructure noise.[104] Finally, they report that ARFIMA models also produce useful 22-day-ahead volatility forecasts in low volatility periods. However, due to the focus and test methodology of their study, they provide no answer to the question of which prediction approach provides the best volatility forecast.[105]

Although the SPA test allows for a simultaneous comparison of multiple forecasts, it provides little information on which particular model is superior to the benchmark model. Further, the SPA test requires the specification of a benchmark model and thus cannot be applied when no natural benchmark exists.[106] Becker and Clements (2008) close this gap by applying the MCS approach developed by Hansen et al. (2003). The objective of this approach is to identify a final set of optimal volatility forecasting models that do not significantly differ with respect to their predictive ability.[107,108]

Becker and Clements (2008) analyse different approaches for the prediction of S&P 500 index volatility. In their analysis, they compare the forecasting performance of implied volatility, time series models, and combined forecasts. They measure the implied volatility of S&P 500 index options using the VIX. Due to the time series characteristics of the series, they fit models from the GARCH, stochastic volatility, ARMA and ARFIMA classes to the data. Their sample comprises the period from January 1990 to October 2003.[109,110]

Their evaluation results for the individual forecasts indicate that volatility forecasts based on time series models using realised volatility outperform option-implied

[104] They suggest that the poor relative performance of model-free implied volatility can be explained by the existence of a volatility risk premium.
[105] See Martin et al. (2009), pp. 89-101.
[106] See Hansen et al. (2003), p. 841.
[107] See Becker and Clements (2008), p. 123.
[108] The criterion to determine the optimal models is user-specified, e.g., the MSE. See Chapter 6.6.2 for a detailed presentation of the MCS approach.
[109] Actual volatility is estimated based on the sum of squared 30-minute S&P 500 index returns.
[110] See Becker and Clements (2008), pp. 122-124.

volatility forecasts.[111] If the forecast set is extended by combined forecasts, then the combination of an ARMA and an ARFIMA model provides the most accurate predictions when using the MSE as the loss function. Applying the QLIKE loss function, they report that the combination of an ARMA model, an ARFIMA model, a stochastic volatility model and the VIX exhibits the lowest forecast errors.[112]

The results of the MCS approach demonstrate that the individual forecasts of the VIX, the GARCH models (with one exception), and the stochastic volatility models are significantly outperformed by combined forecasts. Moreover, the outcome demonstrates that volatility forecasts produced by the ARMA and ARFIMA models provide essential information for the prediction of S&P 500 index volatility, because the MCS contains no other individual forecasts. Overall, the combination of volatility forecasts from short and long memory models using realised volatility provide the best forecasting results.[113]

5.2.7. Summary

The above-described initial debate in the literature concerning the predictive power of implied volatility for US stock market volatility largely presents evidence that BS implied volatility provides better volatility forecasts than historical volatility models.[114,115] Further, these studies typically indicate that implied volatility is a biased predictor of stock market volatility.[116] Because this finding suggests that implied volatility might not contain all information on future volatility or that the

[111] This result holds for the MSE and quasi-likelihood (QLIKE) loss functions.
[112] See Becker and Clements (2008), pp. 129-131.
[113] See ibid., p. 132.
[114] See Latané and Rendleman (1976), Schmalensee and Trippi (1978), Chiras and Manaster (1978), Beckers (1981), Christensen and Prabhala (1998), Ederington and Guan (2005), and Fleming et al. (1995).
[115] Poon and Granger (2003) provide a comprehensive literature review and summary that implied volatility tends to be more appropriate for predicting volatility than historical volatility models. See Poon and Granger (2003), pp. 506-507.
[116] Additionally, Szakmary et al. (2003), who examine futures options data for different asset classes over a wide range of 35 markets, also report that implied volatility is a biased estimator of realised volatility. See Szakmary et al. (2003), p. 2173.

information is not processed correctly by the option pricing model employed,[117] chapters 5.2.2 and 5.2.3 cover studies that investigate whether this can be attributed to measurement errors in implied volatility and/or realised volatility.

With respect to measurement errors in implied volatility due to market imperfections, research published by Christensen and Prabhala (1998) and Ederington and Guan (2002) suggests that the EIV problem can at least partially explain the forecasting bias of implied volatility. By accounting for the EIV problem, these papers provide evidence that implied volatility principally has strong predictive power and generally subsumes the information contained in historical volatility.

Considering the effects of using intraday returns to measure ex-post volatility, Poteshman (2000) demonstrates that the forecasting bias of implied volatility is reduced and its forecasting performance improves when realised volatility is computed based on intraday returns. Blair et al. (2001) also report that intraday returns are better measures of ex-post volatility and provide better volatility forecasts than daily returns. Despite these findings, they document that implied volatility is more informative than historical volatility including intraday returns.

Analysing whether the misspecification of an option pricing model induces forecasting bias for implied volatility, the empirical studies by Poteshman (2000), Shu and Zhang (2003), and Chernov (2007) described in Chapter 5.2.4 do not permit drawing a clear conclusion. While Poteshman (2000) and Chernov (2007) report that replacing the BS model with the Heston model to derive implied volatility reduces forecasting bias, Shu and Zhang (2003) present contradictory results. Moreover, Shu and Zhang (2003) do not report significant differences in the forecasting performance of implied volatility from the BS model and the Heston model. Additionally, they demonstrate that implied volatility outperforms historical volatility and subsumes the information contained in historical volatility models.

[117] As mentioned in Chapter 4.1.2, a test of the information efficiency of options markets cannot be separated from testing the hypothesis of whether the employed option pricing model is correct. Thus, a rejection of the hypothesis can also imply that the option pricing models employed do not hold.

This finding is confirmed by Jiang and Tian (2005) and Taylor et al. (2010), who in addition to BS implied volatility, use model-free implied volatility to forecast stock market volatility. With respect to the relative forecasting performance of model-free implied volatility and BS implied volatility, Jiang and Tian (2005) document that model-free implied volatility provides better prediction results than BS implied volatility. In contrast, Taylor et al. (2010) find that BS ATM implied volatility generally outperforms model-free volatility.

In addition to the development and application of more suitable option pricing models, the papers by Li (2002), Martens and Zein (2004), and Becker et al. (2006) presented in Chapter 5.2.5 report that long memory models provide good volatility forecasts that can improve implied volatility forecasts by incorporating incremental information. As a consequence, they suggest combined volatility forecasts based on implied volatility and long memory models.

The evidence provided by the above papers using encompassing regressions is supplemented in Chapter 5.2.6 by the results of studies that employ loss functions to assess the performance of different volatility forecasting methods. While Fung and Hsieh (1991) report that implied volatility contributes little additional information on future volatility relative to simple rolling volatility forecasts, Gospodinov (2002) finds that implied volatility contains valuable information on future volatility and suggests combining individual forecasts from implied volatility and historical volatility to improve forecasting performance. By extending the set of time series models to ARFIMA models based on realised volatility, Koopman et al. (2005), Becker and Clements (2008), and Martin et al. (2009) find that long memory models provide useful and occasionally better prediction results than historical volatility models. In accordance with Li (2002), Martens and Zein (2004), and Becker et al. (2006), they suggest that combining individual volatility forecasts from different prediction approaches can improve the performance of volatility forecasts. Having presented a literature overview of empirical studies on the prediction of US stock market volatility, the following Section addresses the empirical results for the German stock market.

5.3. Empirical Results for the DAX Options Market

Relative to the US stock market, there is less empirical research on the forecasting performance of implied-volatility and return-based models for the German stock market. Bluhm and Yu (2001) represent an early contribution and compare the predictive power of BS implied volatility with the historical mean, the exponentially weighted average, four ARCH-type models, and a stochastic volatility model using DAX returns and DAX options data from July 1996 to June 1999.[118] These models are applied to compute one-trading-day-ahead, ten-trading-days-ahead, 45-calendar-days-ahead, and 180-trading-days-ahead DAX volatility forecasts. Thereafter, they employ different error measures to evaluate the prediction results.[119] Using the mean absolute per cent error (MAPE) to evaluate 45-calendar-days-ahead and 180-trading-days-ahead DAX volatility forecasts, they find that implied volatility and the stochastic volatility model provide the best forecasting results. Furthermore, their findings demonstrate that implied volatility outperforms ARCH-type models for both forecast horizons. However, they find that ARCH-type models produce useful DAX volatility forecasts over shorter horizons.[120] Overall, they report that the model ranking depends on the forecast horizon and the employed error measure, such that their results do not permit identifying a superior overall prediction method for DAX volatility forecasts.[121]

Claessen and Mittnik (2002) conduct an analysis similar to that of Bluhm and Yu (2001). They also investigate whether the VDAX provides better DAX volatility

[118] Bluhm and Yu (2001) use the VDAX to measure the implied volatility of DAX options.
[119] See ibid., pp. 1-6.
[120] They evaluate short-term DAX volatility forecasts based on the linear-exponential (LINEX) loss function and boundary violations that are applied in the Value-at-Risk approach.
[121] See ibid., pp. 12-18.

forecasts than return-based volatility models, including several GARCH models.[122] In contrast to Bluhm and Yu (2001), they also examine the performance of combined volatility forecasts. Moreover, they apply an extended model set, namely the historical moving average model, a random walk model, a standard and a modified GARCH(1,1) model, an autoregressive model for squared returns, and an extended GARCH(1,1) model where implied volatility is added as an explanatory variable, to produce DAX volatility predictions. The forecasting performance is compared for different forecast horizons based on different evaluation criteria for the period from February 1992 to December 1995.[123,124]

Within the class of individual forecasts, they report that the GARCH(1,1) model extended by implied volatility provides the best prediction results in terms of the MSE across different forecast horizons, whereas implied volatility exhibits high prediction errors. With respect to combined forecasts, they find that the combination of volatility forecasts based on implied volatility and the GARCH(1,1) model provides better forecasting results and outperforms the extended GARCH(1,1) model. They suggest that it is possible to correct the forecasting bias of implied volatility by including implied volatility in the GARCH(1,1) model or using combined forecasts. Overall, they conclude that past DAX returns provide no additional information relative to DAX implied volatility.[125]

Lazarov (2004) extends the above-cited research on forecasting DAX index volatility by estimating time series volatility models based on realised volatility.[126] Further, he applies realised volatility to compare the forecasting performance of several volatility prediction methods, including forecasts based on option-implied volatility.[127] In addition to using implied volatility, he considers the GARCH(1,1) model, a GARCH

[122] In addition to VDAX data, they use transaction data on short-term near ATM DAX index options.
[123] They apply the MAE, the MSE, and the proportion of correctly predicted directions to evaluate the DAX volatility predictions.
[124] See Claessen and Mittnik (2002), pp. 302-309.
[125] See ibid., pp. 312-320.
[126] Additionally, he incorporates realised volatility into time series volatility models to capture additional information on future volatility.
[127] Lazarov (2004) computes realised volatility as the sum of squared 5-minute DAX returns.

model for realised variance, an ARFIMA model for realised variance, an extended version of these time series volatility models where implied volatility is included, and an ad-hoc linear regression model that accounts for the volatility risk premium.[128] The data set contains transaction data on DAX futures and options from January 1999 to July 2002.[129]

Lazarov's (2004) out-of-sample forecast evaluation results demonstrate that the GARCH(1,1) model provides poor prediction results and that the model is outperformed by option-implied volatility and the ARFIMA model. Moreover, he reports that the extended GARCH model exhibits a better performance than the standard GARCH(1,1) model. However, the prediction results of both GARCH models are outperformed by the ARFIMA model based on realised variance. Interestingly, the ad-hoc linear regression model provides prediction results similar to those of the ARFIMA model. In summary, he concludes that implied and realised variance capture the same information for predicting DAX index volatility.[130]

While Lazarov (2004) considers short-term DAX volatility forecasts of up to ten days, Raunig (2006) examines whether DAX volatility is predictable over longer horizons. In his study, he computes daily DAX volatility forecasts over 10, 20, and 45 trading days for the period from December 1997 to July 2005. In contrast to previous findings, he demonstrates that DAX volatility is predictable up to 40 trading days. Because the predictability test results indicate that more sophisticated volatility models are useful for the prediction of DAX volatility over longer horizons, he performs an out-of-sample forecasting experiment to investigate the performance of GARCH models and option-implied volatility. In particular, he compares the predictive ability of the GARCH(1,1) model, the Glosten-Jagannathan-Runkle GARCH (GJR-GARCH)(1,1) model, and the VDAX.[131]

[128]The ad-hoc model represents a regression of realised variance on lagged implied volatility.
[129]See Lazarov (2004), pp. 41-47
[130]See ibid., pp. 57-64.
[131]See Raunig (2006), pp. 363-364.

The forecasting performance is evaluated based on different loss functions, namely MAE, MSE, and HMSE, where latent volatility is measured as the sum of daily squared returns over the forecast horizon. The evaluation results for the selected prediction methods are mixed. On the one hand, the MAE and MSE loss functions suggest, that with one exception, DAX volatility forecasts based on the VDAX outperform the GARCH models. On the other hand, the HMSE criterion indicates that the GJR-GARCH model provides the lowest prediction errors. Because the evaluation method applied by Raunig (2006) is based on the sum of daily squared returns, the mixed results are might be induced by the choice of a noisy volatility measure and should be interpreted with caution.[132]

To close the gap in the empirical literature on the information content of option prices on the future distribution of the underlying asset, Wilkens and Röder (2006) investigate whether higher moments of option-implied distributions contain information on the underlying's future moments. Their study is based on a transaction data set of DAX options from January 1999 to December 2000.[133] By applying the BS model, the Gram/Charlier density expansion model, and two models with mixtures of lognormal distributions, they extract the entire risk-neutral distribution from DAX option prices and examine the information content of option-implied volatility, skewness and kurtosis on future moments. In addition to analysing option-implied moments, they also consider the forecasting ability of historical higher moments.[134]

Using univariate regressions, they find that option-implied volatilities principally contain information on future volatility and offer greater explanatory power for future volatility than does historical volatility.[135] The results of the encompassing regressions suggest that the explanatory power of implied volatility does not increase when historical volatility is added to the regression model. Thus, DAX

[132]See Raunig (2006), pp. 370-371.
[133]For further empirical results regarding Euro-Bund-Future options that are not presented here, see Wilkens and Röder (2006).
[134]See ibid., pp. 50-53.
[135]In particular, they report that volatility forecasts based on implied volatility from the BS model outperform alternative option pricing models in terms of the adjusted coefficient of determination.

implied volatilities subsume the information on future volatility provided by historical volatility. Further, they report that volatility forecasts based on implied volatility from more complex option pricing models do not deliver better prediction results than the BS model. Additionally, they also document that neither implied nor historical skewness or kurtosis provide information on the corresponding future moments.[136]

Muzzioli (2010) applies the promising approach developed by Jiang and Tian (2005) that uses model-free implied volatility to forecast DAX index volatility. Muzzioli (2010) compares the forecasting performance of BS implied volatility, model-free implied volatility, and volatility forecasts based on time series models. Her sample comprises intraday data on DAX index options from January 2001 to December 2006. Further, she avoids the non-synchronous data problem by matching prices in a one-minute interval. Following Andersen and Bollerslev (1998), she measures ex-post volatility as the sum of 5-minute squared DAX returns.[137]

Estimating encompassing regressions, she finds that DAX volatility forecasts based on implied volatility subsume the information provided by historical returns, regardless of whether she uses an AR(1) or a GARCH(1,1) model to predict DAX volatility. Furthermore, she reports that BS implied volatility covers the information contained in model-free implied volatility. According to Muzzioli (2010), the inferior results of model-free implied volatility are might be due to measurement errors induced by illiquid options that are used to calculate model-free implied volatility based on a cross section of option prices.[138] Moreover, to investigate whether the forecasting performance of model-free implied volatility is affected by the implementation method selected, she employs different implementation methods with respect to the extrapolation of the strike price domain.[139] Her results indicate that the imple-

[136] See Wilkens and Röder (2006), pp. 64-68.

[137] See Muzzioli (2010), pp. 561-563.

[138] In comparison to ATM options, the liquidity of options near the minimum, respectively maximum, available strike is typically lower.

[139] The methodology of her basic implementation method follows Jiang and Tian (2005), who assume a constant volatility function beyond the maximum and minimum strike price. In addition to Jiang and Tian (2005), she considers two different implementation methods. First, as an alternative to the assumption of a constant volatility function beyond the available strike

mentation method suggested by Jiang and Tian (2005) provides better results than alternative computation methodologies.[140]

Similar to Muzzioli (2010), Schöne (2010) examines the information content of BS implied volatility measured by the VDAX and model-free implied volatility provided by the DAX-Volatilitätsindex New (VDAX-New). While Muzzioli (2010) focuses on the comparison of DAX volatility forecasts based on implied volatility and time series models, Schöne (2010) investigates whether different ex-post volatility measures affect the evaluation of DAX volatility predictions. In addition to the classical close-to-close estimator that is based on squared daily returns, he employs four different alternative ex-post volatility measures using additional information such as the day's high, low, and opening price.[141] Additionally, he applies a high-frequency volatility measure that is based on the daily sum of squared 15-minute DAX returns. The study considers daily DAX, VDAX, and VDAX-New data from January 1992 to October 2009 and DAX intraday data from January 2001 to October 2009.[142]

The volatility predictions are evaluated by performing univariate regressions in which ex-post DAX volatility (measured by different daily and high-frequency volatility measures) is regressed on DAX volatility forecasts based on the VDAX, respectively the VDAX-New. In general, the regression results suggest that the VDAX-New provides more information on future DAX volatility than the VDAX. However, Schöne (2010) reports that the degree of the information content of each volatility index on future DAX volatility depends on the applied ex-post volatility measure. With respect to the higher information content of the VDAX-New, he argues that this can be explained by the shorter maturity of the underlying DAX options (the VDAX is computed from DAX options with 45 days to maturity, whereas the VDAX-New

prices, she uses an extrapolation method that matches the slope of the smile at the minimum, respectively maximum, strike value. Second, she does not extend the set of strike prices beyond the available range.

[140]See Muzzioli (2010), pp. 571-575.

[141]In particular, he applies the Parkinson estimator, the Garman/Klass estimator, the Rogers/Satchell estimator, and the Yang/Zhang estimator. See Schöne (2010) for a brief description of the estimators.

[142]See ibid., pp. 625-645.

is calculated from options that expire in 30 days).[143] As it uses different forecast horizons, the evaluation approach applied by Schöne (2010) does not allow for a proper comparison of the predictive ability of the VDAX and the VDAX-New.

Recently, Tallau (2011) also analyses whether the VDAX and/or the VDAX-New provide information for the prediction of DAX volatility. In particular, he compares the predictive power of both volatility indices, the asymmetric GJR-GARCH(1,1) model, and the RiskMetrics approach based on univariate and encompassing regressions.[144] He uses daily data on the VDAX, the VDAX-New, and the DAX for the period from January 1992 to December 2008 and considers different forecast horizons.[145] Despite their different computational methodologies, he documents several empirical similarities for both volatility indices across different forecast horizons. First, the VDAX and VDAX-New provide a biased estimate of future DAX volatility. Second, they subsume the information that is contained in historical returns. However, the regression results demonstrate that the VDAX-New provides more information on future volatility than the VDAX and time series models. Further, in contrast to Muzzioli (2010), he finds that the VDAX-New comprises the information provided by the VDAX. Because Tallau (2011) uses an ex-post volatility measure based on squared daily returns and Muzzioli (2010) employs a high-frequency volatility measure, the conflicting results may be due to Tallau's (2011) application of a noisy volatility measure.[146,147]

Overall, the above studies present evidence that DAX implied volatility contains helpful information for the prediction of DAX volatility. While the articles by Bluhm and Yu (2001) and Raunig (2006) report mixed results regarding the forecast ranking, recent studies by Lazarov (2004), Wilkens and Röder (2006), Muzzioli (2010), and Tallau (2011) suggest that DAX implied volatility provides better volatility forecasts than time series models based on historical returns. In addition, Claessen and

[143] See ibid., pp. 645-654.
[144] The RiskMetrics framework was developed by J.P. Morgan and is based on the exponentially weighted moving average method. See Morgan (1996).
[145] See Tallau (2011), pp. 47-50.
[146] Tallau (2011) uses the sum of squared daily returns over 30 days to measure ex-post volatility.
[147] See ibid., pp. 59-72.

Mittnik (2002) indicate that combined volatility forecasts using the information from implied volatility and historical returns are a reasonable complement to individual forecasts. Further, Lazarov (2004) presents the promising result that the forecasting performance of ARFIMA models is similar to that of DAX implied volatility. With respect to the forecast ranking of BS implied volatility and model-free volatility, Muzzioli (2010) documents that BS implied volatility covers the information contained in model-free volatility, whereas Tallau (2011) presents the opposite results.

5.4. Model Selection

As the literature review presented above demonstrates that a comprehensive comparison of forecasting approaches has yet to be performed for the German stock market, this thesis is designed to help close this gap. After the general description of volatility forecasting models in Chapter 4 and the literature review of their empirical forecasting performance presented in this Chapter, this Section explains the selection of the specific models used in this thesis to forecast DAX volatility. In addition to the general model characteristics and the empirical results presented above, model selection is also based on the models' ability to match the observed DAX IVS as documented in Chapter 3. Because the advantages and disadvantages of each forecasting approach are deeply discussed in Chapter 4, this Section draws on this discussion and focuses on particular aspects.

Chapter 4 introduced four different classes of option pricing models: local volatility models, the concept of model-free implied volatility, stochastic volatility models, and jump-diffusion models. While local volatility models are in principle able to perfectly replicate the observed DAX volatility smiles, they exhibit weaknesses when the dynamics of the complete DAX IVS have to be generated.[148] Britten-Jones and Neuberger (2000) developed the concept of model-free implied volatility to overcome these deficiencies. Moreover, Muzzioli (2010) and Tallau (2011) report evidence that

[148]See Wallmeier (2003).

DAX volatility forecasts based on this concept subsume the information contained in historical returns. Therefore, model-free implied volatility is used in this study to produce DAX volatility forecasts.

Although stochastic volatility models provide a more comprehensive approach than local volatility models to explain the characteristics of DAX implied volatilities, they are not considered in this study, because the inclusion of additional risk factors makes the models more complex and reliable parameter estimation more difficult.[149,150] Further, jump-diffusion models are also not applied in this thesis due to their inability to generate certain patterns of the DAX volatility term structure.[151]

Finally, despite its well-documented weaknesses, the BS model is employed as the benchmark model in this study, because it can be regarded as a heuristic rule applied by many market participants.[152] Thus, in the following Chapter, the BS model and model-free implied volatility are used to derive implied volatilities for the prediction of DAX volatility.

Moreover, time series models are employed to produce DAX volatility forecasts. First, GARCH models based on daily DAX returns are employed due to their ability to reproduce volatility clustering effects that are documented in Chapter 3. Second, the above-cited promising empirical results concerning the forecasting performance of long memory models motivated the selection of the ARFIMA and HAR models.[153,154] In addition to individual forecasts, this study also considers combined forecasts, because forecast combinations have been found to outperform individual forecasting models in many areas.[155]

[149]See ibid., p. 239.
[150]In addition, Bates (1996b) and Das and Sundaram (1999) also report that unreasonably high parameters are necessary to reproduce the pronounced volatility smiles of short-term options.
[151]See Das and Sundaram (1999).
[152]In particular, this study analyses, i. a., whether model-free implied volatility based on a broader information set provides better forecasting results than standard BS implied volatility.
[153]See, e.g., Martens and Zein (2004) and Becker et al. (2006).
[154]An analysis of whether the model features are conform to the empirical characteristics of the data is presented in the next Chapter.
[155]See, e.g., the comprehensive studies by Makridakis and Hibon (2000), Stock and Watson (1999), and Marcellino (2004).

6. Forecasting DAX Volatility

This Chapter focuses on the generation and evaluation of DAX volatility forecasts. After analysing the characteristics of the DAX return and volatility series, various volatility forecasting models are estimated. Then, information criteria are used to select the appropriate model specifications. Because the data set contains long time series that cover clearly different volatility periods, a number of structural break tests are performed. The next Section describes the forecasting method and evaluation approach used in this study. Finally, the prediction results for different forecast horizons are presented, evaluated and compared with previous findings in the literature.

6.1. Data Description

While the empirical analysis of the stylised facts of implied DAX volatilities in Chapter 3 considers the complete IVS, this Chapter examines DAX volatility forecasts that require different data filters. Specifically, the calculation of volatility forecasts based on implied DAX volatilities uses cross-sectional data (option prices) and not the complete IVS. Thus, the change in research topic requires an adjustment of the underlying data filters.

First, DAX ITM options and DAX options with a remaining lifetime of fewer (more) than 5 (360) days are excluded from the sample because of liquidity concerns.[1] An

[1]See Aït-Sahalia and Lo (1998), who argue that ITM options are not completely reliable due to infrequent trading. See Aït-Sahalia and Lo (1998), p. 517.

DAX ITM option is defined as a call (put) option with a moneyness level below (above) 0.97 (1.03).[2] Second, to ensure that the daily estimated IVS is as stable as possible within the day, only DAX option contracts that are traded between 4 p.m. and 5 p.m. are considered.[3] The reduction of the sample to one trading hour is primarily induced by the calculation of model-free volatility, which is based on a complete range of option prices.[4] Jiang and Tian (2005) and Muzzioli (2010) also employ model-free volatility to calculate volatility forecasts and use option contracts from a particular trading hour. The trading hour between 4 p.m. and 5 p.m. is selected, as most trading activities typically occur before the trading phase ends at 5:30 p.m. Finally, this study excludes all DAX options with implied volatilities below 5% and above 120% to eliminate the effects of outliers.[5] After all filters have been applied, the remaining data set contains 0.81 million option contracts. The following Section presents some descriptive statistics, time series plots, correlograms and a correlation matrix for the return and volatility series investigated in this study.

6.2. Descriptive Statistics

To present some descriptive statistics for daily DAX model-free volatilities and DAX BS ATM implied volatilities, both implied volatility series are computed based on the selected options data set. As the construction of the implied volatility series was described in the previous Chapters, I refer to the corresponding passages.[6] The

[2] Jiang and Tian (2005) employ a similar filter.

[3] Note that a minimum set of option prices is necessary to ensure a precise estimate of the IVS.

[4] The reduction of the sample size can be explained as follows: First, the implementation of model-free volatility requires the conversion of the observed option prices to implied volatilities. Then, smoothing methods are applied to the implied volatilities to obtain a complete set of implied volatilities. These smoothed implied volatilities are reinserted into the BS equation to derive the desired range of option prices. The crucial point is that the calculation of BS option prices is based on the index level. If the above time interval is not restricted, variation in the index level (e.g., over one day) can induce substantial pricing errors due to non-synchronous data.

[5] See also Chapter 3.2.3.

[6] The BS implied volatilities are calculated according to Section 2.3. Then, the Nadaraya-Watson estimator is used to compute daily BS implied volatilities for a fixed grid of maturities and moneyness levels (see Section 3.1.3). The optimal bandwidths for the Nadaraya-Watson estimator

implied volatilities of both series are determined for a DAX option with constant maturity (here: one month) on an annualised basis. The choice of maturity is motivated by the VDAX-NEW, which also has a maturity of one month. In addition to the implied volatility series, daily realised DAX volatilities are estimated as the cumulative sum of squared 5-minute DAX returns over one day.[7] Finally, DAX returns are calculated from daily DAX index levels at 5 p.m. The descriptive statistics are reported in Table 6.1.[8]

Table 6.1.: Descriptive statistics of volatility and return series from 2002 to 2009

	mfv	*lnmfv*	*bsatm*	*lnbsatm*	*rvola*	*lnrvola*	*rdax*
Mean	0.250	-1.470	0.243	-1.515	0.218	-1.691	0.000
Std. Dev.	0.112	0.402	0.118	0.431	0.144	0.556	0.016
Skewness	1.248	0.526	1.405	0.523	2.260	0.424	-0.112
Exc. Kurt	0.877	-0.542	1.689	-0.445	7.704	-0.102	4.541
Minimum	0.116	-2.155	0.103	-2.277	0.037	-3.306	-0.080
Maximum	0.684	-0.380	0.812	-0.208	1.302	0.264	0.102

Source: own calculations.

Table 6.1 indicates that daily DAX returns also match the well-known stylised facts that are typically observed for financial returns: negative skewness and excess kurtosis. Similar to the related literature, the table reports that model-free volatility and BS ATM implied volatility, are on average, greater than realised volatility.[9] Chernov (2007) attributes this result to the existence of a volatility risk premium that causes implied volatility to exceed realised volatility.[10] Further, DAX options, on average, exhibit a slightly greater model-free volatility than BS ATM implied volatility. Muzzioli (2010) suggests that this can be explained by the fact that model-free volatility also includes information from non-ATM options that overall have higher

are determined by minimising the penalisation function (3.22). The daily model-free volatilities are derived as shown in Section 4.1.2.

[7]See Section 6.6.1 for a detailed description of estimating realised DAX volatilities from DAX intraday returns. An explanation of how 5-minute DAX returns are computed from DAX futures prices is given in Chapter 3.2.2. Analogous to both implied volatility series, the realised volatility is expressed on an annual basis.

[8]In the following tables, DAX BS ATM implied volatility is abbreviated *bsatm*, DAX model-free volatility *mfv*, realised DAX volatility *rvola*, and DAX return *rdax*. The prefix *ln* together with the abbreviation for the times series denotes the corresponding log-series.

[9]Muzzioli (2010) reports the same finding. See Muzzioli (2010), p. 568.

[10]See Chernov (2007), p. 420.

overall implied volatilities than ATM options.[11] Moreover, all three linear volatility series are positively skewed and exhibit positive excess kurtosis.

The time series plot of daily DAX returns in Figure 6.1 displays two extremely volatile market periods that are described in Section 3.2.4. During both periods, the DAX dramatically declined and volatility increased considerably. Furthermore, the plot depicts volatility clustering for the DAX return series, which is typical of many financial return series. Figure 6.2 indicates that each of the two implied DAX volatility series is closely related to realised DAX volatility. It also illustrates that realised DAX volatility is more "spiky" than either implied volatility series. The correlation matrix of the volatility series, which is provided in Table 6.2, confirms this close relationship between implied and realised DAX volatility. It also suggests that DAX model-free and DAX BS ATM implied volatility are highly correlated. The correlation between each volatility series and the DAX return is negative, which indicates the existence of the leverage effect.[12]

Table 6.2.: Correlation matrix

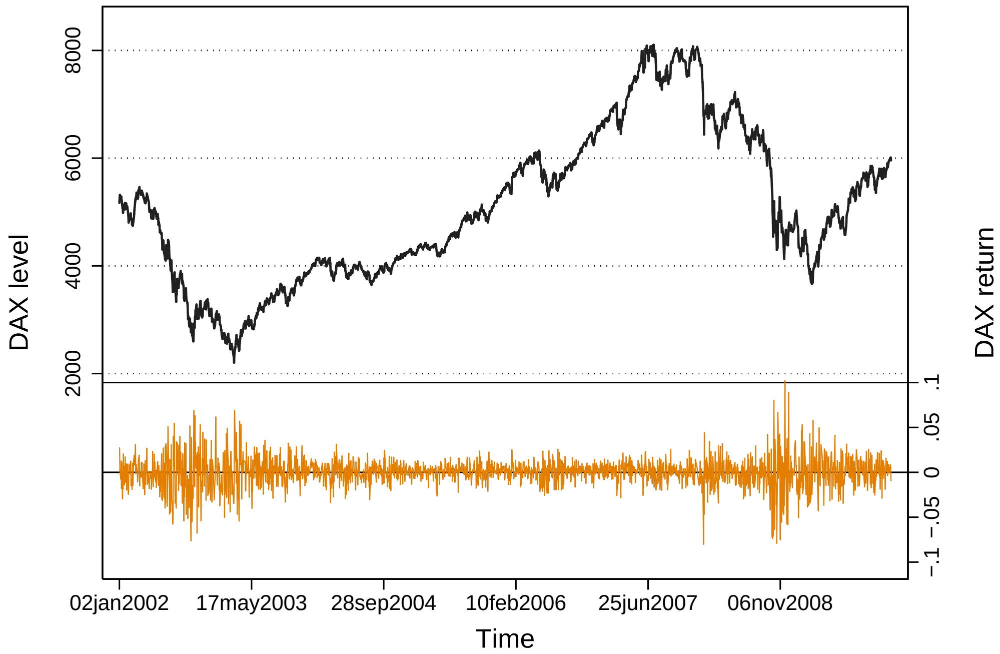
	mfv	*bsatm*	*rvola*	*rdax*
mfv	1.000			
bsatm	0.995	1.000		
rvolva	0.888	0.907	1.000	
rdax	-0.115	-0.120	-0.138	1.000

Source: own calculations.

To study the autocorrelation structure, Figure 6.3 contains a correlogram for each volatility and return series. While the sample autocorrelations for DAX returns are not generally significantly different from zero, the correlograms for all three volatility series indicate significant positive serial correlations. Moreover, the slow decay of the

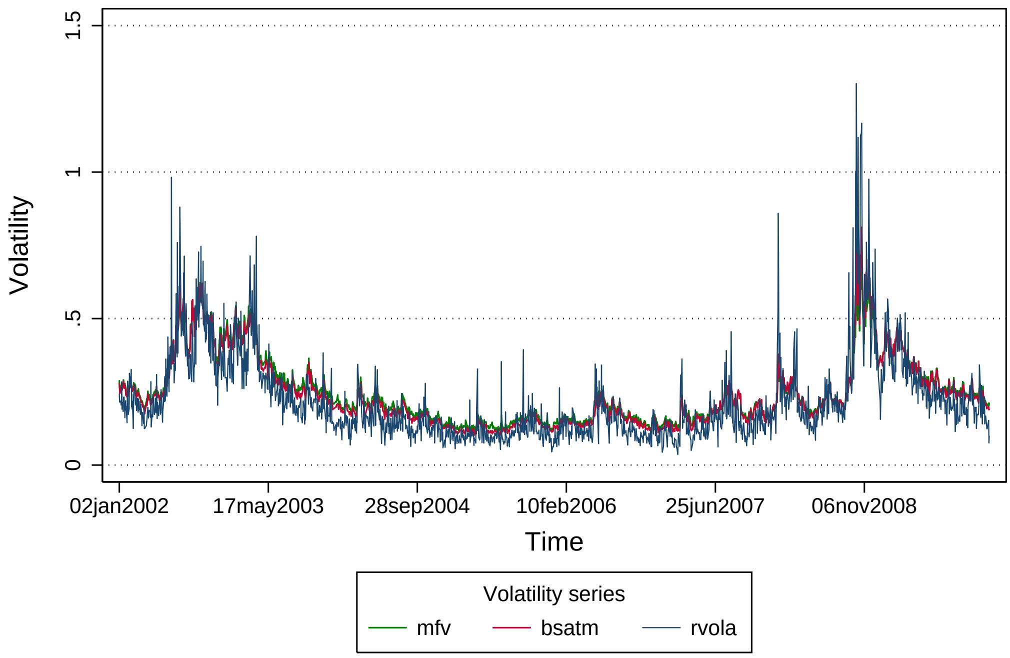
[11] See Muzzioli (2010), p. 569.

[12] See Fengler (2012), who also reports a negative correlation between DAX returns and BS 1M ATM implied volatility for the period from 2000 to 2008. See Fengler (2012), p. 123.



Source: EUREX, own calculations.
Note: DAX level (left axis) and DAX return (right axis).

Figure 6.1.: DAX level and DAX return series



Source: EUREX, own calculations.
Note: Daily implied volatilities refer to DAX options with 1 month to maturity.

Figure 6.2.: DAX volatility series

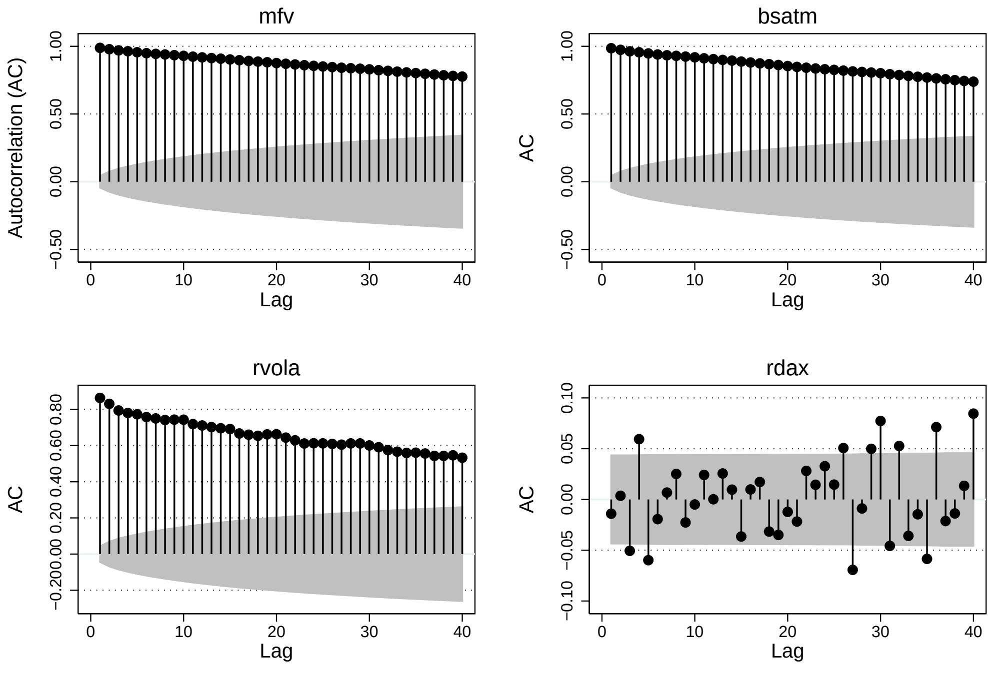


Figure 6.3.: Correlograms of DAX volatilities and return series

sample autocorrelations in the volatility series provides evidence for the existence of long memory effects.[13]

6.3. Tests Results for Unit Roots, Long Memory, and ARCH Effects

After presenting some descriptive statistics, this Section reports the results of various unit-root tests to investigate whether the time series are stationary. Several statistical tests have been developed to test for stationarity. This study employs the augmented Dickey-Fuller (ADF) test, the Dickey-Fuller generalised least-

[13]Koopman et al. (2005) obtain related results for daily S&P 100 returns, realised volatilities, and implied volatilities for the period from January 1997 to November 2003. See Koopman et al. (2005), p. 452.

squares (DFGLS) test suggested by Elliott et al. (1996), and the KPSS test developed by Kwiatkowski et al. (1992).

The ADF test is widely used and based on the null hypothesis that the series is non-stationary. The ADF test can be computed for different cases that differ with respect to whether, under the null hypothesis, the process contains a drift and whether the test regression contains a constant or trend term. Given the time series plot of DAX returns in Figure 6.1, the null hypothesis of a random walk without drift is tested against the alternative of a stationary process with a constant, but no trend term. In addition, the plots of the volatility series in Figure 6.2 suggest that the hypothesis of a random walk with drift should be tested against the alternative of a stationary process with a constant, but no trend term.

Further, the DFGLS test suggested by Elliott et al. (1996) is implemented, as Elliott et al. (1996) report that the test has a significantly higher power than the original ADF test. The DFGLS test represents a modified Dickey-Full test in which a possible drift and/or deterministic linear trend is removed from the time series via a generalised least squares regression before the unit root test is performed.[14,15] The null hypothesis of the DFGLS test states that the series follows a random walk, whereby a drift is possible. Finally, the KPSS test is selected to complement the unit root tests because, in contrast to the ADF and the DFGLS tests, it uses the null hypothesis of a stationary time series.

Table 6.3 below summarises the findings of the above unit root tests. While none of the unit root tests suggest the existence of a unit root for the DAX return series, the results for the volatility series are mixed. Specifically, the ADF test and the DFGLS test reject the non-stationarity hypothesis for realised volatility at the 5% level, but the KPSS test indicates non-stationarity. For the implied volatility series, the null hypothesis of a random walk without drift is accepted by the ADF test in all cases at the 5% significance level. These findings are consistent with the results

[14]See Elliott et al. (1996), pp. 815-826.

[15]Here, the DFGLS test is performed without the assumption of a linear trend. The highest lag order is selected based on the method proposed by Schwert (1989).

of the KPSS test. In contrast, with one exception (log model-free volatility), the DFGLS test provides evidence in favour of a stationary implied volatility series. The ADF test of the null hypothesis of a random walk with drift is rejected for both implied volatility series at the 5% level.[16] The literature argues that such contradictory results can be induced by the inability of the previous unit root tests to distinguish between the integration orders one and zero.[17] Thus, in the following, the Geweke-Porter-Hudak (GPH) test and a modified version of Robinson (1995b) (denoted ROB here) are performed to detect long memory effects in the series.[18]

To estimate the long memory parameter d of a fractionally integrated process of the form $(1 - L)^d X_t = \varepsilon_t$ where ε_t is stationary with zero mean, Geweke and Porter-Hudak (1983) suggest a semiparametric procedure. The GPH test has the appealing feature that it avoids the need to specify the ARMA structure.[19] Further, a modified version of the GPH estimator proposed by Robinson (1995b) is employed. Hidalgo and Robinson (1996) argue that under certain conditions, this modified test statistic is likely more efficient.[20] The GPH and ROB estimates of the long memory parameter are presented for typical values of the tuning exponent θ in the last section of Table 6.3.[21] While the null hypothesis that d is equal to zero is rejected for all volatility series at the 1% level, the long memory parameter is not significantly different from zero for DAX returns at any conventional significance level. An additional t-test indicates that d is not significantly different from one at the 5% level for the implied volatility series, which provides evidence that the series are non-stationary. For the realised volatility series, the t-test indicates that the long memory parameter is significantly different from one at the 1% level. In summary,

[16] For the results, see Table B.1 in the Appendix.
[17] See Kirchgässner and Wolters (2007), p. 180.
[18] For instance Coakley et al. (2011), McAleer and Medeiros (2008), and Gospodinov et al. (2006) apply the GPH test. Andersen et al. (2003) and Ashley and Patterson (2010) use the adjusted version of the GPH test provided by Robinson (1995b).
[19] See Baum (2000), p. 41.
[20] See Hidalgo and Robinson (1996), p. 173.
[21] The tuning exponent indirectly specifies the number of ordinates that enter the log-periodogram regression (see Baum (2000), p. 39). A typical value of θ for the GPH estimate is 0.5 and 0.8 for the Robinson estimate. See Ashley and Patterson (2010), p. 68.

Table 6.3.: Results of stationarity, long memory, and ARCH LM tests

test	*mfv*	*lnmfv*	*bsatm*	*lnbsatm*	*rvola*	*lnrvola*	*rdax*
ADF (10)	-2.46	-2.30	-2.63	-2.38	-3.76	-3.48	-13.64
p-value	12.6%	17.2%	8.6%	14.6%	0.3%	0.8%	0.0%
ADF (20)	-2.37	-2.15	-2.63	-2.28	-3.23	-2.83	-10.59
p-value	15.1%	22.4%	8.6%	17.7%	1.8%	5.3%	0.0%
DFGLS	-2.28	-1.91	-2.44	-2.00	-3.01	-2.15	-8.20
t_c	-1.953	-1.954	-1.949	-1.953	-1.946	-1.947	-1.946
KPSS	4.20	4.74	3.73	4.48	2.81	3.86	0.17
t_c	0.463	0.463	0.463	0.463	0.463	0.463	0.463
GPH ($\theta = 0.5$)							
d	0.96	0.94	0.93	0.91	0.71	0.75	0.14
t	10.13	9.94	9.41	9.76	7.01	8.95	1.17
p-value	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	24.7%
GPH ($\theta = 0.8$)							
d	0.87	0.86	0.86	0.86	0.55	0.54	-0.02
t	28.33	27.81	27.61	27.47	16.66	17.08	-0.79
p-value	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	42.9%
ROB ($\theta = 0.5$)							
d	0.96	0.94	0.93	0.91	0.71	0.75	0.13
t	10.14	9.94	9.41	9.76	7.01	8.95	1.04
p-value	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	30.2%
ROB ($\theta = 0.8$)							
d	0.85	0.85	0.84	0.84	0.53	0.52	-0.02
t	28.30	27.78	27.59	27.42	16.59	17.01	-0.76
p-value	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	44.6%
ARCH LM							
t	1,838.4	1,901.3	1,764.0	1,885.7	1,028.4	1,191.4	472.0
p-value	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%

Source: own calculations.

these findings suggest that long memory effects are responsible for the mixed results of the unit root tests.

Volatility clustering represents another stylised fact of financial returns that is also observed for DAX returns (see Figure 6.1). Engle (1982) suggests a Lagrange multiplier test (ARCH LM) to test for the presence of volatility clustering or ARCH effects. The null hypothesis of this test states that the time series is free of ARCH effects. The results of the ARCH LM test presented in Table 6.3 indicate that the null hypothesis is rejected at the 1% level for each series. This suggests the presence of ARCH effects in all series.[22] Further, the sample autocorrelation function (ACF) of the squared DAX return residuals, which is depicted in Figure 6.4, demonstrates that all sample autocorrelations are significantly different from zero at the 5% level.[23,24] In addition, the combined hypothesis that all sample autocorrelations in the squared DAX return residuals are zero is tested based on the Portmanteau test (LB hereafter) developed by Ljung and Box (1978). The results of the LB test indicate that the combined hypothesis can clearly be rejected at the 1% level when the lag length is 10 or 20.[25]

In summary, these findings and the observed excess kurtosis in DAX returns indicate that ARCH effects exist in DAX return series considered here.[26] As GARCH models are able to capture these effects, various GARCH and EGARCH models are estimated below. In addition to these models, further members of the GARCH family were estimated (e.g., the extended GARCH model, where realised variance is added to the variance equation as an explanatory variable). They are not further considered in this study due to estimation problems or poor prediction results. Moreover,

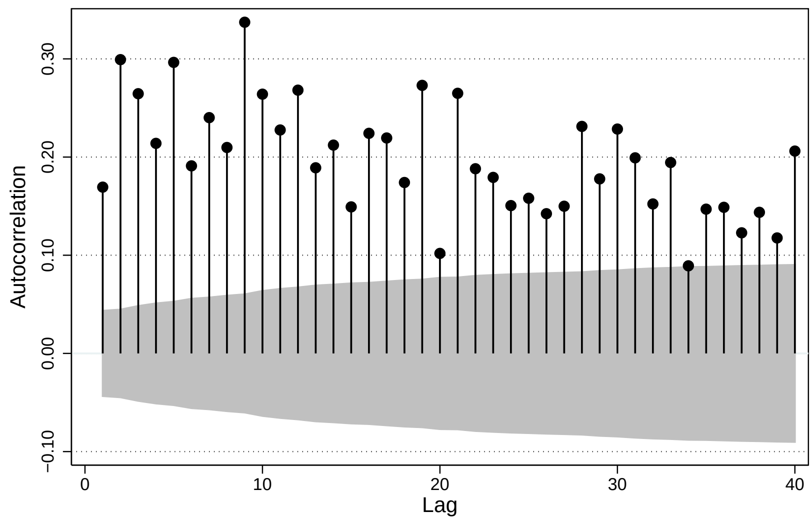
[22] The lag length of the ARCH LM test is set to 10 to capture the features of a GARCH(1,1) process, which can be approximated by an ARCH(q) process when q is sufficiently large.

[23] Based on simulation results, Bollerslev (1988) shows that an analysis of the sample correlation structure of the squared residuals can provide helpful information for model identification issues. See Bollerslev (1988), p. 130.

[24] The squared DAX return residuals are computed based on the mean model, which is described in Section 6.4.

[25] The LB test statistic takes the value 1314.16 (2217.59) when the lag length is 10 (20).

[26] As mentioned above, the ARCH LM test also implies that the volatility series exhibit ARCH effects.



Source: EUREX, own calculations.
Note: Bartlett's formula for MA(q) 95% confidence bands.

Figure 6.4.: Correlogram of squared DAX return residuals

ARFIMA and HAR models are fitted to the volatility series, as the results of the GPH test and the modified version developed by Robinson (1995b) provide evidence that the volatility series are fractionally integrated.

6.4. Identification, Estimation, and Selection of Volatility Time Series Models

This Section addresses the identification, estimation, and selection of the GARCH, ARFIMA, and HAR models employed in this study to forecast DAX volatility.[27] After specifying of the conditional mean equation, the Section discusses the identification of various GARCH-type models. Then, the GARCH models are estimated and information criteria are used to select the "best" GARCH models for DAX volatility prediction. Similarly, this Section presents the identification, estimation, and selection of long memory models to conduct DAX volatility forecasts.

[27]See Chapter 4.2 for a brief introduction to each volatility model.

6.4.1. GARCH Models

Specification of the Conditional Mean

Specifying the variance equation for the GARCH models requires the identification of an appropriate model for the conditional mean. To do so, the sample ACF, the sample partial autocorrelation function (PACF), and the corrected LB test developed Diebold (1988) are computed. The sample ACF of the DAX returns is plotted in Figure 6.5. Diebold (1988) demonstrates that Bartlett's confidence bands are overly conservative in the presence of ARCH effects. It particular, the use of Barlett's original formula can lead to a misspecification of the ARMA process for the conditional mean, since the original formula underestimates the variances of the sample autocorrelations. Thus, the corrected 95% confidence interval $B(\tau)$ for the sample autocorrelation $\widehat{\rho}(\tau)$ is calculated based on Diebold's formula:

$$B(\tau) = 0 \pm 1.96\; S(\tau)^{0.5} \tag{6.1}$$

with

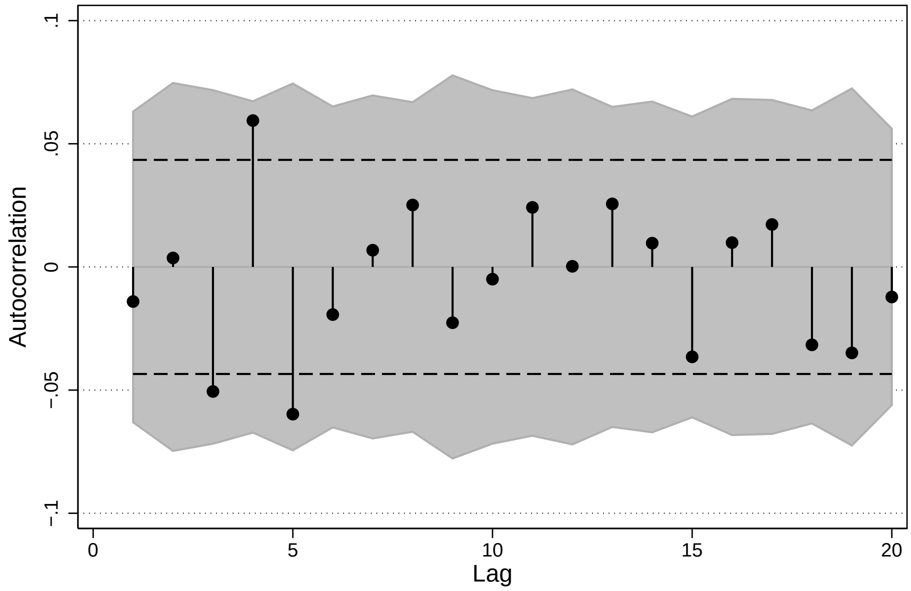
$$S(\tau) = \frac{1}{T}\left(1 + \frac{\widehat{\gamma}(\tau)}{\widehat{\sigma}^4}\right) \tag{6.2}$$

where $S(\tau)$ denotes the variance of the sample autocorrelations, $\widehat{\rho}(\tau)$, $\widehat{\gamma}(\tau)$ is the estimated autocovariance of the squared DAX returns, and $\widehat{\sigma}^4$ represents the squared sample variance of the DAX returns.[28,29] Figure 6.5 depicts the sample autocorrelations of the DAX returns and the corrected confidence intervals. To visualise the effect of the corrected confidence bands, the standard confidence bands based on Bartlett's formula are also depicted, as dashed black lines, in Figure 6.5.

Figure 6.5 indicates that the sample autocorrelations of the DAX returns up to lag 20 fall within the corrected confidence bands. This implies that the DAX returns are not serially correlated. The corrected LB developed by Diebold (1988) is performed

[28] See Diebold (1988), p. 21.

[29] See Krämer and Runde (1994) and Kokoszka and Politis (2008), who show that Barlett's formula should not be used for ARCH processes. Hsieh (1989), for example, provides an application of Diebold's formula.



Source: EUREX, own calculations.
Note: Diebold's (1988) formula for heteroskedastic−corrected 95% confidence bands.

Figure 6.5.: Correlogram of the DAX return series

to test the combined hypothesis that the first k autocorrelation coefficients of the DAX returns are zero. The test statistic of the corrected LB test is given by

$$Q^*(k) = T(T+2)\sum_{\tau=1}^{k}\left(\frac{\sigma^4}{\sigma^4+\gamma(\tau)}\right)\frac{\widehat{\rho}^2(\tau)}{T-\tau}.^{30} \tag{6.3}$$

The p-values of Diebold's corrected LB test can be observed in Figure 6.6. The findings confirm that the DAX returns are not serially correlated up to lag 20. Therefore, based on these results, the conditional mean equation is specified as

$$E[r_t|I_{t-1}] = \mu \tag{6.4}$$

where I_{t-1} denotes the information set at time t.[31]

[30] Diebold (1988) suggests correcting for the original Ljung-Box test statistic, as the empirical test size of the Ljung-Box test is larger then its nominal size if ARCH effects are present. See Diebold (1988), p. 28.

[31] The sample PACF of the DAX returns indicates that some partial autocorrelations are significant at the 5% level. Because the magnitude of these partial autocorrelations is low and the pattern is not systematic, the above specification of the conditional mean is selected. See Figure B.1 in the Appendix.

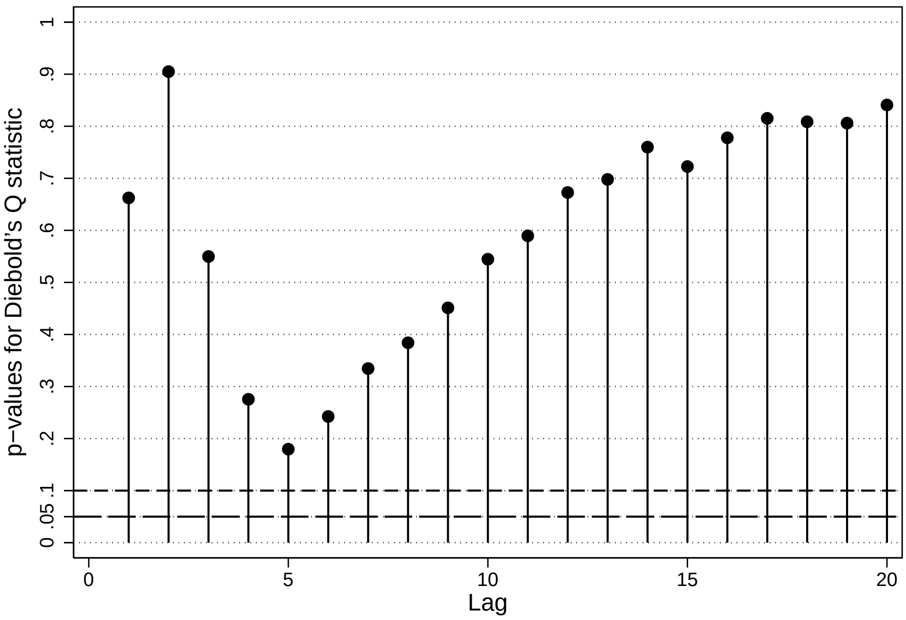


Figure 6.6.: p-values for Diebold's ARCH-robust Q*-statistic

GARCH Model Estimation and Selection

Having specified the mean equation, this Section describes the choice of the error distribution and the determination of the ARCH order p and GARCH order q. Furthermore, this Section also presents the in-sample estimation results and the selection of the GARCH models used in this study to forecast DAX volatility.[32]

Since the pioneering work of Engle (1982), numerous extensions of the basic GARCH model structure have been developed to capture various empirical characteristics, such as asymmetries in the variance or fractional volatility.[33] While a variety of GARCH models exist, in practice the basic GARCH(1,1) model is generally applied. In a comprehensive study, Hansen and Lunde (2005) compare 330 ARCH-type models and report that the forecasting performance of the GARCH(1,1) model is rarely surpassed.[34] For instance, Claessen and Mittnik (2002), Sapusek (2004), Lazarov

[32]Here, the term "GARCH models" not only refers to the standard GARCH model, but also includes asymmetric GARCH models, such as the EGARCH model.

[33]See the "Glossary to ARCH" provided in Bollerslev (2008), which contains an enumeration of over 80 different GARCH models.

[34]See Hansen and Lunde (2005), pp. 881-886.

(2004), Raunig (2006), and Muzzioli (2010) use the GARCH(1,1) model to predict DAX volatility. Thus, the GARCH(1,1) model is a natural candidate for predicting DAX volatility.

As GARCH models assuming a normal error distribution are often unable to completely reproduce the leptokurtosis of financial time series, non-normal error distributions have been suggested to model the observed excess kurtosis.[35] While Bollerslev (1987), Baillie and Bollerslev (1989), and Hsieh (1989), among others, suggest Student's t-distribution, Nelson (1991) and Xu and Taylor (1995) employ the generalised error distribution (GED). Moreover, the choice of error distribution affects forecasting performance. Angelidis et al. (2004) and Wilhelmsson (2006) find that GARCH models with leptokurtic error distributions provide better forecasting results than models with normal error distributions. Thus, this study estimates a GARCH(1,1) model with a leptokurtic error distribution. The GED is selected from the class of leptokurtic distributions, as it is more flexible than Student's t-distribution.[36] Given the large number of observations, precise estimates of the additional parameters should be feasible.

The in-sample estimation results of the GARCH(1,1) model fitted to the DAX return series from January 2002 to December 2009 are presented in Table 6.4. The table demonstrates that the estimates of the ARCH parameter α_1 and the GARCH parameter β_1 are significantly different from zero at the 1% level. Moreover, the use of a non-normal distribution is justified by the value of the estimated shape parameter of the GED, which is significantly lower than two.[37] Similar to some research findings on financial returns,[38] the sum of the estimated ARCH and GARCH parameters is close to one (0.995), which indicates that DAX volatility is highly persistent.[39] Franke et al. (2004) report a related result for daily DAX index returns

[35]See Bollerslev et al. (1992), p. 11.
[36]The variance of the GED depends on two parameters (one for scale and one for shape), whereas the variance of Student's t-distribution is determined by one parameter (the degrees of freedom).
[37]If the shape parameter of the GED is lower than two, the GED is leptokurtic.
[38]See Bollerslev et al. (1992), pp. 14-15.
[39]As noted in Section 4.2.1, the GARCH(1,1) model is stationary for $\alpha_1 + \beta_1 < 1$.

Table 6.4.: Estimation results for GARCH models

	GARCH(1,1)	GARCH(1,2)	EGARCH(2,1)
μ	0.001***	0.001***	4.8E-04**
	(2.2E-04)	(2.2E-04)	(2.3E-04)
ω	1.3E-06***	8.3E-07***	-0.134***
	(5.1E-07)	(3.1E-07)	(0.032)
α_1	0.078***	0.043***	
	(0.012)	(0.009)	
β_1	0.917***	1.509***	0.985***
	(0.011)	(0.080)	(0.004)
β_2		-0.555***	
		(0.076)	
θ_1			-0.202***
			(0.047)
θ_2			0.096**
			(0.046)
γ_1			-0.163***
			(0.063)
γ_2			0.283***
			(0.062)
shape	1.513	1.527	1.623
	(0.051)	(0.100)	(0.092)
ln L	5,986.9	5,991.1	6,030.6
AIC	-11,963.8	-11,970.2	-12,045.3
SIC	-11,935.8	-11,936.5	-12,000.3

Source: own calculations.
Note: Standard error in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. The "shape" parameter refers to the GED distribution of the model errors.

from January 1998 to December 2007.[40] The impact of this outcome remains a disputed matter in the literature.

To capture the persistence of conditional variances, Engle and Bollerslev (1986) suggest the Integrated GARCH (IGARCH) model, which imposes a precise unit root in the autoregressive polynomial.[41] Lamoureux and Lastrapes (1990) advise against the application of IGARCH models to long time series. They demonstrate that high persistence in the variance can be attributed to deterministic shifts in the

[40] See Franke et al. (2011), p. 321.
[41] In particular, $\alpha_1 + ... + \alpha_q + \beta_1 + ... + \beta_p = 1$.

unconditional variance that are not accounted for the model. In contrast, Brooks (2008) notes that the theoretical motivation for non-stationarity in the variance is weak, as non-stationarity in the variance implies that the unconditional variance is unbounded.[42] Because Figure 6.2 provides additional indications that DAX volatility is mean-reverting rather than unbounded, the standard GARCH(1,1) model is used to predict DAX volatility. Lamoureux and Lastrapes (1990) argument that volatility persistence may be due to structural breaks is investigated separately in this study.

In addition to the GARCH(1,1) model, various GARCH models of up to order $p = q = 2$ are estimated to identify the model that provides the "best fit" to the DAX return series. The selection of an appropriate GARCH model is typically based on information criteria, such as the Akaike information criterion (AIC) and the Schwartz information criterion (SIC).[43] Although information criteria have been widely used, a careful interpretation of the results is necessary, as little is known about the statistical properties of the information criteria in the presence of ARCH effects.[44] For this reason, Brooks and Burke (1998) suggest the use of modified information criteria that are also applicable when volatility clustering is observed. Specifically, they propose a modified version of the SIC, termed the heteroskedastic Schwartz information criterion (HSIC), which is based on the Kullback-Leibler discrepancy. They define the HSIC as

$$HSIC \equiv \sum_{t=1}^{T} \log(\widehat{h}_t^2) + g \log(T) \tag{6.5}$$

where $\widehat{h}_t^2$ denotes the estimated conditional variance using the corresponding GARCH model and $g = p + q + 2$ is the total number of estimated parameters. As various criteria have been suggested and no clear favourite can be observed in the literature,

[42]See Brooks (2008), p. 37.
[43]See Zivot (2009), who argues that the use of the classical information criteria is justified by the possibility of expressing GARCH models as ARMA models of squared residuals. See Zivot (2009), p. 126.
[44]See Bollerslev et al. (1994) and Leeb and Pötscher (2009).

Table 6.5.: Diagnostic test results for GARCH models

	GARCH(1,1)	GARCH(1,2)	EGARCH(2,1)
Q1(10)	17.9	17.3	16.1
	5.6%	6.8%	9.8%
Q1(20)	24.2	23.4	23.1
	23.5%	26.8%	28.5%
Q2(10)	11.2	7.4	6.6
	34.4%	68.3%	76.1%
Q2(20)	25.7	22.3	21.9
	17.5%	32.3%	34.6%
ARCH LM(10)	9.8	6.3	7.0
	46.1%	78.6%	72.5%
ARCH LM(20)	24.0	21.1	22.4
	24.3%	39.2%	31.9%
Sign bias	0.5	0.5	-1.0
	61.6%	62.3%	33.1%
Negative sign bias	0.8	0.4	0.0
	39.9%	72.5%	96.8%
Positive sign bias	-2.6	-2.4	-0.3
	1.1%	1.5%	73.9%
Joint effect	10.4	10.1	1.3
	1.5%	1.8%	72.7%

Source: own calculations.
Note: For each test, the first row contains the value of the test statistic and the second row presents the p-value.

all three information criteria are calculated in this study to determine the ARCH order p and the GARCH order q.

Table B.2 in the Appendix provides the results of the information criteria for the various GARCH models. Of the standard GARCH models, the GARCH(1,2) model exhibits the lowest values for each of the three information criteria. Thus, based on these findings, this model is selected to predict DAX volatility. The in-sample estimation results of the GARCH(1,2) model are reported in Table 6.4. Similar to the GARCH(1,1), the estimated ARCH and GARCH parameters are significantly different from zero at the 1% level and the sum of the parameters is again near one.[45] Further, the estimated shape parameter of the GED is less than two.

To investigate whether the two GARCH models completely capture the ARCH effects, the standardised residuals of the GARCH(1,1) and GARCH(1,2) models are analysed. The distributions of the standardised residuals of the GARCH models are examined in a first step. Theory implies that the distributions of the standardised residuals of the estimated GARCH models should approximate the normal distribution.[46] In the following, the Jarque-Bera (JB) test is used to determine whether the skewness and kurtosis of the standardised residuals are equal to zero and three, respectively, which would be their values under normality.[47] Although the values of the JB test statistic for the standardised residuals are much smaller than those for DAX returns,[48] the null hypothesis of normality is nevertheless rejected at all conventional significance levels.[49]

However, the results of the LB test (denoted *Q2*) for the squared standardised residuals of the GARCH models reported in Table 6.5 indicate that the squared standardised residuals no longer exhibit any significant sample autocorrelations. The results of Engle's LM test for ARCH effects, which are also presented in Table 6.5, support this finding for both GARCH models. Additionally, the null hypothesis of zero autocorrelation between the standardised residuals of the GARCH models is assessed based on the LB test. The corresponding test statistic is denoted *Q1* and the results are also presented in Table 6.5. The null hypothesis, which states that there is zero autocorrelation between the standardised residuals, is rejected for both GARCH models. Therefore, while both GARCH model are unable to completely capture the excess kurtosis and skewness of the DAX returns, the results of the LB

[45] Because the sum of the estimated ARCH and GARCH parameters is lower than one, the GARCH(1,2) model is covariance stationary.
[46] See Zivot (2009), p. 127.
[47] The JB test statistic computed for the distribution of the standardised residuals of the GARCH(1,1) (GARCH(1,2)) model is 207.96 (172.21).
[48] The JB test statistic calculated for DAX returns is 1748.61.
[49] The null hypothesis is rejected due to the remaining excess kurtosis and the skewness in the distribution of the standardised residuals.

test and Engle's ARCH LM test indicate that the models are able to reproduce the ARCH effects.

As skewed return distributions can be induced by the leverage effect,[50] four tests are performed to investigate whether asymmetries exist in the standardised residual series of the two GARCH models. If asymmetric effects are detected in the residual series, the application of an asymmetric GARCH is reasonable, as a standard GARCH model does not capture these effects.

The four diagnostic tests were proposed by Engle and Ng (1993) and use the following regression

$$v_t^2 = c_0 + c_1 I_{\varepsilon_{t-1}<0} + c_2 I_{\varepsilon_{t-1}<0}\, \varepsilon_{t-1} + c_3 I_{\varepsilon_{t-1}\geq 0}\, \varepsilon_{t-1} + u_t \tag{6.6}$$

where v_t denotes the standardised residuals of the GARCH models and I is an indicator function that takes the value 1 if the respective condition, e.g., $\varepsilon_{t-1} < 0$, holds and zero otherwise. Based on this regression, the null hypotheses $H_0^i:\ c_i = 0$ (for $i = 1, 2, 3$) and $H_0^j:\ c_1 = c_2 = c_3 = 0$ are tested. The test of the null hypothesis H_0^1 is called the *sign bias test*, the test of the second null hypothesis H_0^2 is called *negative size bias test*, and the test of the third null hypothesis H_0^3 is the *positive size bias test*.[51,52] All four tests are applied to the standardised residuals of the GARCH(1,1) and GARCH(1,2) models. The test results are presented in Table 6.5. Null hypotheses H_0^3 and H_0^j are rejected at the 5% level for the GARCH(1,1) and GARCH(1,2) models. This indicates the presence of asymmetry in the conditional volatility.

To address such asymmetries, several extensions of the standard GARCH model have been suggested. For instance, Nelson (1991) proposes the EGARCH model, Glosten et al. (1993) develop the GJR-GARCH model, Engle and Ng (1993) suggest the Asymmetric GARCH (AGARCH) model, and Zakoian (1994) advises the use

[50] See Bouchaud and Potters (2001), p. 65.
[51] See Engle and Ng (1993), pp. 1757-1763.
[52] In Table 6.5, the test of the combined hypothesis H_0^j is denoted *joint effect*.

of the Threshold GARCH (TGARCH) model. This study employs the EGARCH model to reproduce the asymmetric effects, as the above-cited study by Awartani and Corradi (2005) reports that the EGARCH model exhibits better out-of-sample performance than the GJR-GARCH, TGARCH, or AGARCH model. Furthermore, the EGARCH model has the convenient feature that the conditional variance is always positive. Following Nelson (1991), the EGARCH model with GED residuals is used in this work, as the unconditional variance does not necessarily exist for other leptokurtic distributions (e.g., Student's t-distribution).[53] Similar to the standard GARCH model, the appropriate order of the EGARCH model is determined by information criteria. For this purpose, various EGARCH models are estimated up to order $p = q = 2$. The results of the information criteria for these models are presented in Table B.2 in the Appendix.

While the AIC recommends the use of the EGARCH(2,2) model, the SIC and the HSIC favour the EGARCH(2,1) model. Because the AIC is known to exhibit a tendency towards over-parameterised models, the SIC and HSIC are used to select the EGARCH(2,1) model.[54] The in-sample estimation results of the EGARCH(2,1) model, which are presented in Table 6.4, reveal that the estimates of the model parameters are significantly different from zero at the 5% level.[55] The findings of Engle's ARCH LM test reported in Table 6.5 demonstrate that the EGARCH(2,1) model is able to capture the observed ARCH effects of the DAX return series. Further, Engle and Ng's (1993) sign bias tests indicates that the EGARCH model fully captures the asymmetric behaviour of the conditional volatility, as none of the four null hypotheses are rejected at the 5% level (see Table 6.5). In summary, drawing on these insights, the GARCH(1,1), the GARCH(1,2), and the EGARCH(2,1) are able to reproduce the observed stylised facts of DAX returns. The estimation and selection results for the ARFIMA and HAR models are presented below.

[53]See Nelson (1991), p. 352.

[54]Specifically, the HSIC is more useful in the presence of ARCH effects than the traditional information criteria (see the discussion above).

[55]Moreover, the EGARCH(2,1) model is covariance stationary, as β_1 is lower than one (see Section 4.2.1 for the stationarity condition of an EGARCH model).

6.4.2. ARFIMA and HAR Models

Estimation and Selection

Using the ARFIMA and the HAR model to fit realised DAX volatilities is motivated by the results of the GPH test and its modified version suggested by Robinson (1995b). Both tests indicate that realised DAX volatilities are fractionally integrated (see Table 6.3). Additionally, Figure 6.3 demonstrates that the sample ACF of realised DAX volatilities decays hyperbolically, which is characteristic of fractionally integrated processes.[56] Therefore, various long memory models are fitted to the realised DAX volatility series.[57] In contrast, the linear decay of the sample ACF for implied DAX volatilities provides evidence that the implied volatility series are integrated of order one (see Figure 6.3). This finding agrees with the results of the unit root tests presented in Table 6.3, which generally suggest that the implied volatility series are non-stationary. Therefore, the long memory models are not applied to implied DAX volatilities.

To select an appropriate model, various ARFIMA models of up to order $p = q = 2$ are fitted to realised DAX volatilities. To capture the so-called *weekend effect*, the ARFIMA model equations are extended by a Monday dummy variable.[58,59] Model selection is performed based on the AIC and SIC, as suggested by Crato and Ray (196) for ARFIMA models.[60] The ARFIMA(1,d,1) model is selected from the

[56] See Zivot and Wang (2008), p. 273.

[57] Here, the term "long memory models" also refers to models that are only able to mimic long memory effects, such as the HAR model.

[58] In the following, the coefficient on the Monday dummy variable is denoted by θ.

[59] The weekend effect states that higher stock market return volatility is observed after a weekend. For instance, Harvey and Whaley (1992), who consider S&P 100 index options, demonstrate that implied volatility generally tends to increase on Mondays. They suggest that this variation is due to excess buying pressure induced by traders who are opening their positions on Monday. See Harvey and Whaley (1992), p. 58.

[60] The AIC is defined as $\mathrm{AIC} = -2\log L + 2s$, and the SIC is given by $\mathrm{SIC} = -2\log L + s\log N$, where L denotes the maximum likelihood of the ARFIMA model considered, and $s = 1 + p + q + k + 1$ contains the number of estimated parameters. The number of estimated parameters consists of the ARMA orders p and q, the number of parameters of the mean equation k, and the last summand accounts for the residual variance. See Doornik and Ooms (2006), p. 26.

estimated models, as it exhibits the lowest AIC and SIC values (see Table B.3 in the Appendix).

The in-sample estimation results of the ARFIMA(1,d,1) model are presented in Table 6.6. The output indicates that the estimated ARMA parameters α_1 and β_1, the coefficient of the Monday dummy θ, and the fractional differencing parameter d are significantly different from zero at the 1% level. The estimate of the long memory parameter d suggests that realised DAX volatilities are weakly stationary.[61] To examine whether the estimated model is correctly specified, the residuals of the ARFIMA(1,d,1) model are analysed. The results of the LB test, which are also presented in Table 6.6, demonstrate that the sample autocorrelations of the residuals are not significantly different form zero at the 5% level.[62] Further, the GPH test detects no remaining long memory effects. In summary, the ARFIMA model seems able to capture the dependencies of realised DAX volatilities. Therefore, the ARFIMA model is used to predict DAX volatility.

In addition to the ARFIMA model, the HAR model proposed by Corsi (2009) is also able to reproduce long memory effects. An introduction to the HAR model was presented in Section 4.2.2. As mentioned above, the HAR model does not formally belong to the class of long memory models, but it does also capture a hyperbolic decaying ACF. Similar to the ARFIMA model, the model equation for the HAR model is extended by including a Monday dummy variable to capture the weekend effect. The in-sample estimation results of the HAR model are presented in Table 6.6. The table indicates that the estimated coefficients of the HAR model are all significantly different from zero at the 5% level. As the LB test of the model residuals suggests that the HAR model does not capture all serial dependencies, the sample ACF and PACF of the residuals are examined.[63] Figure B.2 in the Appendix indicates that the sample autocorrelation at lag $k = 2$ of the residual series is significantly different from zero. Thus, the HAR model is modified by adding

[61] A weakly stationary process is characterised by a fractional parameter d below 0.5.
[62] The results of the LB test are denoted $Q1(10)$ and are provided in Table 6.6.
[63] For the results of the LB test, see Table 6.6.

two-period lagged realised DAX volatility to the model equation. The estimated parameters of the modified HAR model are also presented in Table 6.6. The table provides evidence that the parameter estimates are all significantly different from zero at the 5% level. Moreover, the null hypothesis of the LB test cannot be rejected at the 5% significance level after extending the model by adding two-period lagged realised DAX volatility. Finally, the GPH test indicates that the residuals are not fractionally integrated (see Table 6.6). Therefore, the modified HAR model is able to account for the observed dependencies of realised DAX volatilities. The next Section examines the potential effects of structural breaks on the estimation results of the presented forecasting models.

6.5. Structural Breaks

The discussion in Section 3.2.4 suggests that DAX volatility changed considerably over the long sample period (see also Figure 3.3). Further, the estimation results of the standard GARCH models provide evidence that the high persistence of the conditional variance may be due to regime switches.[64] Moreover, the literature notes that long memory effects can arise from the occurrence of structural breaks and regime switching.[65,66] Thus, the following Section investigates whether high persistence in variance and long memory effects can be explained by structural breaks. If these features are the result of structural changes, then the above time series models are misspecified and further considerations are necessary.[67] Before the effects of structural breaks are analysed, the following Section provides a brief overview of selected structural break tests.

[64] Recall the arguments advanced by Lamoureux and Lastrapes (1993) cited above.

[65] Banerjee and Urga (2005) note that long memory effects can be induced by structural breaks. See Banerjee and Urga (2005), p. 22.

[66] Alternatively, Banerjee and Urga (2005) highlight that long memory can be attributed to the aggregation of processes.

[67] For instance, Gourieroux and Jasiak (2001b) argue that the application of a strong fractional model can lead to spurious results if long memory effects are due to infrequent regime switching. See Gourieroux and Jasiak (2001b), p. 38.

Table 6.6.: Estimation results for long memory models

	ARFIMA(1,d,1)	HAR	modified HAR
μ	-1.647***	-0.039**	-0.037**
	(0.306)	(0.019)	(0.019)
α_1	0.974***		
	(0.013)		
β_1	-0.921***		
	(0.036)		
d	0.385***		
	(0.043)		
β^d		0.369***	0.371***
		(0.033)	(0.033)
β^{2d}			0.101***
			(0.030)
β^w		0.416***	0.304***
		(0.045)	(0.057)
β^m		0.185***	0.194***
		(0.036)	(0.035)
θ	-0.045***	-0.065***	-0.069***
	(0.013)	(0.016)	(0.016)
ln L	-27.3	-30.4	-24.2
AIC	66.5	70.7	60.5
SIC	100.2	98.8	94.1
Q1(10)	16.0	26.1	11.9
	9.9%	0.4%	29.4%
GPH	0.1	0.0	0.0
	91.5%	99.3%	99.0%

Source: own calculations.
Note: Standard error in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. The parameter θ denotes the coefficient of the Monday dummy variable and β^{2d} represents the coefficient of the two periods lagged realized volatility.

6.5.1. Testing for Structural Breaks

Chow (1960) develops the classical test for detecting a structural change. Based on the assumption that the breakdate is known, he divides the sample into two subperiods, estimates a model for each subperiod, and finally tests whether the two parameter sets are equal. If the residuals are homoscedastic, the classical F statistic is used to test the equality hypothesis; otherwise the Wald test statistic is applied.[68,69]

Because the breakdate is often not known a priori, Quandt (1960) suggests calculating a sequence of Chow statistics over all possible breakdates and deriving the breakdate using the point with the maximum Chow statistic.[70] While the significance of the test statistics is evaluated based on the critical values of a chi-squared distribution when the breakdate is known a priori, these critical values cannot be used if the breakdate is unknown. Andrews (1993) and Andrews and Ploberger (1994) solve this problem by determining the asymptotic null distributions of the test statistics and providing the corresponding asymptotic critical values. However, the test developed by Andrews (1993) and Andrews and Ploberger (1994), called the sup$F$ test, only considers the occurrence of one structural break. Thus, Bai and Perron (1998) extend their solution by developing a method that allows the researcher to test for multiple structural breaks. In the following, the sup$F$ test developed by Andrews (1993) and the extension developed by Bai and Perron (1998) are presented, as both tests are used in this study.[71] The the sup$F$ test developed by Andrews (1993) is described first.

Assume a linear regression model of the form

$$y_i = \mathbf{x_i}\beta_\mathbf{i} + \varepsilon_i \quad (i = 1, ..., n) \tag{6.7}$$

[68]See Hansen (2001), p. 118.

[69]Recall, that the traditional Chow (1960) test assumes constant variances. See Maddala (2008), p. 391.

[70]Note that the procedure does not consider "all" possible breakdates, but rather subperiods of the full sample that are determined by a trimming parameter. See Hansen (2001), p. 119.

[71]See ibid., pp. 119-121.

where y_i denotes the observed values of the dependent variable at time i, $\mathbf{x_i}$ represents a $k \times 1$ vector of independent variables,[72] and $\beta_\mathbf{i}$ is a $k \times 1$ vector of possibly time-varying coefficients. The null hypothesis that no multiple structural breaks exist is given by

$$H_0: \quad \beta_\mathbf{i} = \beta_\mathbf{0} \quad (i = 1, ..., n) \tag{6.8}$$

and is tested against the alternative that at least one coefficient is time-varying.

When testing for one unknown structural break, the null hypothesis is tested based on the following F statistic

$$F_i = \frac{\hat{\mathbf{u}}^\top \hat{\mathbf{u}} - \hat{\mathbf{u}}(\mathbf{i})^\top \hat{\mathbf{u}}(\mathbf{i})}{\hat{\mathbf{u}}(\mathbf{i})^\top \hat{\mathbf{u}}(\mathbf{i})/(n-2k)} \tag{6.9}$$

where $\hat{\mathbf{u}}$ denotes the residuals from the unsegmented model and $\hat{\mathbf{u}}(\mathbf{i})$ represents the residuals from the segmented model. The segmented model is estimated for two subsamples that are determined by the breakpoint at time i. As the breakpoint is unknown a priori, the calculation of the F statistic is repeated for all possible breakdates $i = n_h, ..., n - n_h \quad (n_h \geq k)$ where $n_h = [nh]$, and h is a trimming parameter. If the supremum of the F statistics, $\sup F = \sup_i F_i$, is excessively large, the null hypothesis of one unknown structural break is rejected.[73,74]

As noted above, Bai and Perron (1998) extend this approach to multiple breaks. Under the assumption that there are m breakpoints, the minimal residual sum of squares of $m + 1$ segmented linear regressions can be written as

$$RSS(i_1, ..., i_m) = \sum_{j=1}^{m+1} rss(i_{j-1} + 1, i_j) \tag{6.10}$$

[72] To add an intercept to the regression equation, the first component of $\mathbf{x_i}$ is set equal to one.
[73] See Zeileis et al. (2003), pp. 110-111.
[74] The structural break analysis in this study is performed based on the R software package "strucchange" developed by Zeileis et al. (2003). In the software package, the approximate asymptotic p-values of the test statistic are computed using the method developed by Hansen (1997).

where $rss(i_{j-1}+1, i_j)$ denotes the minimal residual sums of squares for the regression of the j-th segment. The breakdates $\hat{i}_1, ..., \hat{i}_m$ are determined by solving the optimisation problem

$$\min_{i_1,...,i_m} RSS(i_1, ..., i_m) \tag{6.11}$$

over all partitions $(i_1, ..., i_m)$ with $i_j - i_{j-1} \geq n_h \geq k$. Bai and Perron (2003) suggest a dynamic programming approach to identify the global minimisers. In this framework, the solution is recursively described by

$$RSS(\Im_{m,n}) = \min_{mn_h \leq i \leq n-n_h} [RSS(\Im_{m-1,i}) + rss(i+1, n)] \tag{6.12}$$

where $\Im_{m,n} = \{i_1, ..., i_m\}$ represents a set of breakpoints.[75] In the next Section, the impact of structural breaks on long memory effects is analysed for realised DAX volatilities using the approach of Choi and Zivot (2007) that builds on the Bai and Perron (1998) test.

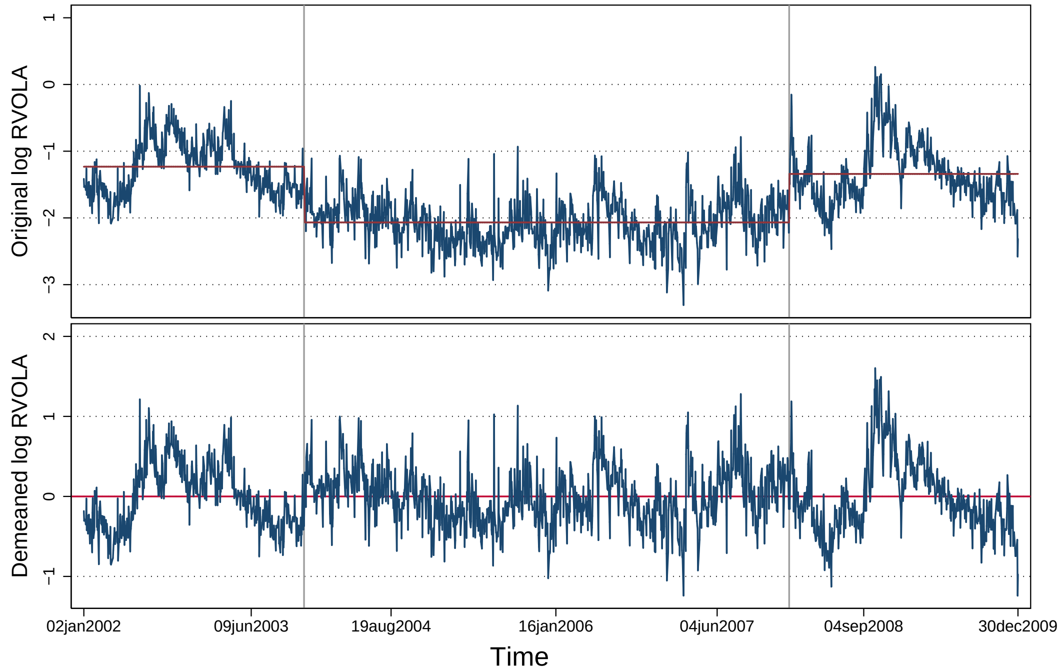
6.5.2. Testing for Long Memory Effects in the Presence of Structural Breaks

As mentioned above, both structural breaks and regime switches can induce long memory effects.[76] As the GPH test and the modified version developed by Robinson (1995b) indicate the presence of long memory effects in realised DAX volatility (see Table 6.3), and the time series plot of the series in Figure 6.2 depicts remarkable changes in realised DAX volatility over the long sample period, the existence of structural breaks and their effect on the long memory property of realised DAX volatilities is analysed in the following.[77]

[75] See Zeileis et al. (2003), p. 112.

[76] See, e.g., Granger and Teräsvirta (1999), Diebold and Inoue (2001), and Granger and Hyung (2004).

[77] This analysis is essential for selecting of the ARFIMA and HAR models. If the long memory effects of realised DAX volatilities can be fully attributed to structural breaks, the detection of structural breaks is more important in predicting DAX volatility than the ability of time series models to capture long memory effects.



Source: EUREX, own calculations.
Note: DAX realized volatility is abbreviated by RVOLA.

Figure 6.7.: DAX log realised volatility with structural breaks in the mean

Choi and Zivot (2007) propose an approach to investigate this issue. They examine the relationship between long memory and structural breaks in the forward discounts of five G7 countries. First, they estimate the long memory parameter under the assumption that no structural break appeared in the prevailing series. Then, they use Bai and Perron's (2003) method to detect multiple structural breaks in the mean of each forward discount series. Based on this result, they calculate demeaned forward discount series. The demeaned series are simply computed as $\hat{u}_t = y_t - \hat{c}_j$ where y_t denotes the original forward discount series and $\hat{c}_j$ represents the estimated mean for each regime of the forward discount series. Next, they re-estimate the long memory parameter for the demeaned series and compare their findings with the previous results. They report that even after the elimination of structural breaks, long memory effects exist in each country's forward discount series. In the following,

Table 6.7.: Results of the Bai and Perron (1998) test

A. Optimal number of breaks				
number of breaks	0	1	2	3
RSS	629.8	490.5	318.2	316.4
SIC	3403.7	2910.0	2044.5	2048.5
B. Breakdates				
break	1	2		
breakdate	2003/11/21	2008/01/15		
C. Regimes				
regime	1st	2nd	3rd	
from	2002/01/02	2003/11/21	2008/01/15	
to	2003/11/20	2008/01/14	2009/12/30	

Source: own calculations.

Choi and Zivot (2007) method is applied to examine whether the long memory effect observed for realised DAX volatilities is due to structural breaks.[78]

First, the Bai and Perron (2003) test is performed to identify multiple structural changes in a linear model. Here, realised DAX volatility is regressed on a constant. Then, the SIC is used to determine the optimal number of breaks.[79,80] The results of the Bai and Perron (1998) test are presented in Table 6.7. They demonstrate that the optimal partition is obtained for two breakpoints. The related breakdates are also reported in Table 6.7. A time series plot of realised DAX volatilities including the estimated breakdates is depicted in the upper panel of Figure 6.7.

[78] In this study, the Choi and Zivot (2007) method is adjusted with respect to the estimation of the long memory parameter: the long memory parameter is not estimated using the modified log-periodogram regression developed by Kim and Phillips (2006); instead an ARFIMA model is fitted to realised DAX volatilities, where the ARFIMA parameters are estimated by maximum likelihood.

[79] Bai and Perron (2003) recommend the use of the SIC, as the AIC tends to overestimate the number of breaks. See Bai and Perron (2003), p. 14.

[80] Generally, the choice of the trimming parameter reflects a trade-off between the need for a minimum sample size to obtain precise parameter estimates and the objective of capturing structural breaks at the beginning and end of the sample period. Further, Bai and Perron (2003) argue that in the presence of serial correlation and/or heterogeneity (including heteroskedasticity) in the data, a higher trimming parameter should be used (see ibid., p. 15). Based on these considerations, the trimming parameter is set to 0.2 in this study. As the test results are plausible and agree with a simple visual examination of the time series, this seems a reasonable choice.

Table 6.8.: Estimation results of the ARFIMA model before and after removing structural breaks

	ARFIMA(1,d,1) (original series)	ARFIMA(1,d,1) (demeaned series)
μ	-1.647***	-0.037
	(0.306)	(0.175)
α_1	0.974***	0.949***
	(0.013)	(0.020)
β_1	-0.921***	-0.866***
	(0.036)	(0.054)
d	0.385***	0.334***
	(0.042)	(0.056)
θ	-0.045***	-0.044***
	(0.013)	(0.013)
Q1(10)	16.01	16.31
	9.9%	9.1%
GPH	0.1	-1.0
	91.5%	33.4%

Source: own calculations.
Note: Standard error in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. The parameter θ denotes the coefficient of the Monday dummy variable.

The detected breakdates can be linked to certain historical events.[81] The first breakdate marks the end of the volatile period at the beginning of the sample, is driven by investors' fears of an impending recession in the US and the Iraq war in 2003. The second breakdate corresponds to the beginning of the financial crisis in 2008. The period between the two volatile market phases reflects a more silent market characterized by a rising DAX. Thus, the findings of the Bai and Perron (1998) test are reasonable, as they correspond to historical events. In the following, the identified breakdates are used to calculate the demeaned realised DAX volatility series.

The demeaned realised DAX volatility series is constructed by removing the mean of each regime from the original realised DAX volatility series. The lower panel of Figure 6.7 provides a time series plot of the demeaned series. In the next step, the ARFIMA(1,d,1) model is estimated for the original and demeaned time series.

[81]See also the discussion in Section 3.2.4 regarding DAX volatility regimes.

The estimation results are presented in Table 6.8. The table demonstrates that the parameters estimated for each ARFIMA model are all significantly different from zero at the 1% level. Further, after the elimination of the structural breaks in the mean, the estimated fractional parameter is below that obtained for the original series but still significant.[82] This suggests that the level of the long memory parameter for the considered realised DAX volatility series is driven by structural breaks. However, the long memory effect does not entirely disappear if structural breaks are removed from the DAX volatility series. Thus, the above ARFIMA model overestimates the long memory effect of realised DAX volatility but nevertheless captures an actual feature of the time series. As the prediction of DAX volatility is based on rolling windows of fixed sample sizes of 500 observations, a series of subsamples that are not affected by the observed structural breaks exists. Therefore, the ARFIMA model is used to produce DAX volatility forecasts. The next Section analyses whether the HAR model and the GARCH reveal structural changes of unknown timing.

6.5.3. Testing for Structural Breaks: Results for the HAR model and the GARCH Models

Due to their model structures, the sup$F$ test for structural changes developed by Andrews (1993) can be directly applied to the modified HAR and the GARCH models. First, the Andrews (1993) test is used to test for parameter instability in the modified HAR model. The test results are presented in Table 6.9.[83] The table provides evidence that the null hypothesis cannot be rejected for the modified HAR model at the 5% level. Thus, the test results suggest that no structural change emerged in the parameters of the modified HAR model across the sample. Next, the findings for the GARCH models are described.

[82] The LB test and the GPH test indicate that the ARFIMA model is able to capture the dependencies of the original and demeaned realised DAX volatility series.

[83] Similar to the previous Section, the trimming parameter is set to 0.2.

Table 6.9.: Results of the Andrews (1993) test

model	sup$F$	p-value
modified HAR	16.74	9.8%
GARCH(1,1)	3.52	45.4%
GARCH(1,2)	3.83	40.2%
EGARCH(2,1)	3.95	38.2%

Source: own calculations.

To test for structural breaks in the variance, Hansen (2012) proposes first regressing the squared residuals of a linear regression model on a constant and, second, applying the Andrews (1993) sup$F$ test to this squared residuals regression.[84] In an ARCH setting, the squared standardised residuals of a GARCH model are regressed on a constant. Then, the Andrews (1993) test is performed to test the null hypothesis that no structural break with an unknown change point appeared in the residual series. Applying this procedure to the selected GARCH models, yields the test results provided in Table 6.9. They indicate that the null hypothesis is not rejected at the 5% level for any of the three GARCH models. Based on these findings, the modified HAR model and the three GARCH models can be directly used to predict DAX volatility without taking structural parameter changes into account. After selecting the volatility forecasting models and conducting the structural break analysis, the forecasting methodology and evaluation approach applied in this study are introduced.

6.6. Volatility Proxy, Evaluation Approach, and Forecasting Methodology

First, this Section explains the selection of the employed volatility proxy, realised volatility. Next, it describes the calculation of realised volatility for the DAX returns series. Thereafter, the Section presents the evaluation approach used to determine whether one or more volatility forecasting models provide superior DAX volatility

[84]See Hansen (2012), p. 30.

predictions. Finally, the forecasting methodology employed in this study is discussed.

6.6.1. Volatility Proxy

As the variable of interest, volatility, is unobservable, a volatility proxy is necessary to evaluate DAX volatility forecasts. While prior studies use squared daily asset returns as a conditionally unbiased estimator for latent variance, the concept of realised variance has recently become increasingly widespread.[85,86] The choice of the volatility proxy is important, as Hansen and Lunde (2006) demonstrate that the application of a conditionally unbiased volatility proxy instead of the true latent variable can lead to a different ranking of volatility forecasts.

Hansen and Lunde (2006) investigate the empirical rankings of volatility prediction models using various volatility proxies. They demonstrate that a ranking of volatility forecasting models based on realised variance is more likely to be consistent with the true ranking, as realised variance is typically a more precise volatility measure than squared daily returns.[87] In addition, Andersen et al. (2006) note that realised volatility is a natural benchmark for the evaluation of volatility forecasts, as it does not rely on an explicit model.[88] Therefore, the basic concept of using realised volatilities as a volatility proxy is presented in the following.

Assume that the logarithmic price process of a given asset follows a continuous semi-martingale $p_t, t \geq 0$ given by

$$p_t = p_0 + \int_o^t \mu_\tau d_\tau + \int_0^t \sigma_\tau dW_\tau \tag{6.13}$$

[85] See Xiao (2013), p. 57.

[86] For instance, Bluhm and Yu (2001) and Claessen and Mittnik (2002) use the standard deviation as an estimate of the average volatility over a fixed time interval, whereas Lazarov (2004), Muzzioli (2010), and Schöne (2010) compute various realised variance estimators based on DAX intraday returns.

[87] See Hansen and Lunde (2006), pp. 98-100.

[88] See Andersen et al. (2006), p. 830.

where W represents standard geometric Brownian motion, $(\mu_\tau)_{\tau\geq 0}$ denotes a finite variation càglàg drift process and $(\sigma_\tau)_{\tau\geq 0}$ is an adapted càdlàg volatility process. Then, the quadratic variation (QV) of the return process is[89,90]

$$r_{0,t} := p_t - p_0 = \int_o^t \mu_\tau d_\tau + \int_0^t \sigma_\tau dW_\tau \tag{6.14}$$

over the interval $[0, t]$ is

$$QV(0,t) = \int_0^t \sigma_\tau^2 d\tau \tag{6.15}$$

with

$$\sum_{j=1}^{n} (p_{j\Delta} - p_{(j-1)\Delta})^2 \xrightarrow{m.s.} QV(0,t) \quad \text{for } n \to \infty.^{91} \tag{6.16}$$

The index variable n denotes the number of high-frequency intervals over $[0, t]$ with length $\Delta = n^{-1}$.

The sample-path variation of the squared return process over the interval $[0, t]$ is also called integrated variance (IVAR). In a continuous-time setting, the quadratic variation equals the integrated variance

$$\text{IVAR}(0,t) := \int_0^t \sigma_\tau^2 d\tau = QV(0,t).^{92} \tag{6.17}$$

As intraday asset returns are available for very small intervals, the IVAR can be estimated over the interval $[0, 1]$ using the standard realised variance

$$RV^n := \sum_{j=1}^{n} (p_{j\Delta} - p_{(j-1)\Delta})^2 := \sum_{j=1}^{n} r_{j\Delta,n}^2 \tag{6.18}$$

where $r_{j\Delta,n}^2$ denotes squared intraday returns.[93]

[89]The quadratic variation reduces to (6.15), as the quadratic variation of the finite drift process μ_t is zero and the quadratic variation of a Wiener process over $[0, t]$ is equal to t.
[90]The application of quadratic variation to measure volatility was first suggested by Andersen and Bollerslev (1998).
[91]See Hautsch (2012), pp. 195-196.
[92]See Andersen and Benzoni (2009), p. 560.
[93]See Hautsch (2012), p. 197.

Formally, realised variance is a consistent estimator of quadratic variation due to the semi-martingale form of the price process and the assumed unlimited sampling frequency. However, in practice, the assumption that the price process is sampled over arbitrarily small intervals is limited by the observed transaction frequency.[94] Further, in a more realistic setting, the price process described by equation (6.13) is affected by a noise component that includes so-called market microstructure effects. Due to the high sampling frequency of DAX futures, the following Section concentrates on the effect of market microstructure noise on the realised variance estimator.

A price process that accounts for these effects is given by

$$p_t = p_t^* + u_t \tag{6.19}$$

where p^* represents a semi-martingale process of the form (6.13) and u_t is a noise component that incorporates market microstructure effects such as price discreteness, bid-ask bounces and non-synchronous trading. Under the assumption that u_t is i.i.d. with zero mean and $E[u_t^2] := \omega^2$, intraday returns behave according to an MA(1) process, and the expected realised variance is given by

$$E[RV^n] = \text{IVAR}(0,1) + 2n\omega^2.^{95} \tag{6.20}$$

The equation demonstrates that in the presence of market microstructure effects, realised variance is a biased estimator of actual variance.[96] Further, for large sampling frequencies, the realised variance estimator is dominated by the noise term, as with $n \to \infty$, the estimator diverges to infinity.[97]

The literature suggests two alternative approaches to reduce the estimation bias caused by market microstructure noise. First, instead of using every tick, realised

[94]See Härdle et al. (2008), p. 277.
[95]See Hautsch (2012), p. 198.
[96]See McAleer and Medeiros (2008), p. 19.
[97]See Hautsch (2012), p. 198.

variance is computed based on lower sampling frequencies, such as 5-, 10-, or 30-minute returns, to avoid (serial) dependencies induced by market microstructure effects.[98] This method is called *sparse sampling*. However, while this procedure leads to reduced estimation bias, the actual variance estimate is less precise, as a reduced information set is considered. Various realised variance estimators have been developed to address this problem, which take market microstructure effects into account but do not exclude (excessive amounts of) information.[99]

For instance, to eliminate serial correlation from high-frequency returns, Andersen et al. (2001b) suggest a two-step approach. First, they fit a moving average model to the intraday return series. Second, the realised variance is estimated by taking the sum of the squared model residuals. Zhang et al. (2005) propose an alternative estimator. They present the so-called *two times scales estimator*, which combines realised variances from various sampling frequencies. Alternatively, Barndorff-Nielsen et al. (2008) recommend a kernel-based estimator of quadratic variation, in which the kernel weights ensure the consistency of the estimator in the presence of market microstructure noise. Bandi and Russell (2006) developed another popular realised variance estimator. Their estimator is based on the notion of selecting an optimal sampling frequency.

In a comprehensive study, Liu et al. (2012) investigate the accuracy of nearly 400 different estimators of asset price variation and find that the 5-minute realised variance estimator nearly always superior. Therefore, this estimator is considered in the following. As the realised variance estimator is only consistent if intraday returns are uncorrelated, the sample autocorrelations of DAX 5-minute returns are examined. If significant serial correlations between the intraday returns are detected, market microstructure noise exists and the estimator is biased. The sample ACF of DAX 5-minute returns is depicted in Figure B.3. The sample autocorrelations at the first and the fifth lags are significantly different form zero at the 5% level, but their mag-

[98]See, for example, Andersen et al. (2000) and Andersen et al. (2003).
[99]See McAleer and Medeiros (2008), p. 21.

nitude is very low. To evaluate their economic relevance, Ebens (1999) suggests a simple approach that is described below.

The approach developed by Ebens (1999) is based on the estimation of a Moving Average (MA) model of order q

$$r_{n,t} = \epsilon_{n,t} + \sum_{i=1}^{q} \psi_{i,t}\epsilon_{n-i,t} \tag{6.21}$$

for high-frequency returns in which the innovations $\epsilon_{n,t}$ are serially uncorrelated.[100] To determine the relationship between the realised variance and the actual variance of the MA(q) process, he derives

$$E\left[\sum_{n=1}^{N} r_{n,t}^2\right] = (1 + \sum_{1}^{q} \psi_{i,t}^2) E\left[\sum_{n=1}^{N} \epsilon_{n,t}^2\right]. \tag{6.22}$$

Thus, if any serial correlation remains in high-frequency returns, the realised variance overestimates the actual variance. The bias can be calculated by plugging the estimated MA coefficients into (6.22). In the following, Ebens's (1999) approach is used to determine the effect of the observed serial correlations on the estimates of DAX volatility.[101]

First, an MA model that includes the moving-average terms for lags one and five is fitted to the DAX 5-minute return series, as the sample autocorrelations at the first and the fifth lags are significantly different from zero at the 5% level (see Figure B.3). The estimation results presented in Table B.4 indicate that for the full sample, the two estimated MA coefficients are significantly different from zero at the 1% level. Next, the estimated MA model parameters are plugged into equation (6.22). The results indicate that the realised variance overestimates the actual volatility by 0.04%. Therefore, the economic effect of the remaining market microstructure noise in 5-minute DAX returns is very low. Thus, the realised variance estimator based on

[100] Ebens's (1999) original specification allows the parameters of the MA model to change over time.

[101] See ibid., p. 11.

5-minute DAX returns is used as volatility measure in this study. A time series plot of the estimated realised 5-minute DAX volatilities for the full sample is provided by Figure 6.7.

6.6.2. Forecast Evaluation

Overview of Forecast Evaluation Techniques

After selecting the volatility proxy, this Section presents the evaluation approach used in this study. The comparison of various forecasting models is typically based on out-of-sample prediction results. The rational is reported various studies obtaining poor out-of-sample forecasting results while providing a good in-sample fit for some models.[102] The literature provides various explanations for this discrepancy. White (2000) argues that in-sample forecast performance is more affected by outliers and data mining-induced overfitting than out-of-sample performance. Enders (2004) notes that the best in-sample model fit will not always provide the best forecasting results due to increasing parameter uncertainty in complex models.[103] Moreover, Poon and Granger (2003) highlight that the design of out-of-sample tests is closer to reality than that of in-sample tests. Therefore, most studies examining the predictive ability of volatility forecasting models use out-of-sample volatility forecasts.[104,105]

The previous Section, demonstrates that several structural breaks occur in the realised DAX volatility series that may have different effects of the results of in-sample and out-of-sample performance tests. Clark and McCracken (2005) analyse the effects of structural breaks on predictive ability tests. They find that structural breaks can explain the differences in the findings of in-sample and out-of-sample performance tests. In particular, they report that out-of-sample performance tests have

[102] See Klein (1992), Fildes and Makridakis (1995), Poon and Granger (2003), and McCracken and West (2004), among others.

[103] See Enders (2004), 82.

[104] See, for example, Koopman et al. (2005), Becker et al. (2006), and Martin et al. (2009).

[105] However, Inoue and Kilian (2005) question this practice. They report that neither data mining nor parameter instability can explain these differences; instead they can be attributed to the higher power of in-sample tests of predictive ability.

the power needed to identify the predictive ability of forecasting models at the end of the sample, which is of particular interest for prediction purposes.[106] As the focus of this study is forecasting DAX volatility, the predictive ability of the selected volatility forecasting models is evaluated based on out-of-sample results.[107] Thus, various out-of-sample forecast evaluation methods are presented below. The discussion begins with an overview of the basic forecast evaluation approaches suggested in the literature. The following discussion is based on the classification proposed by West (2006) that was presented at the beginning of Chapter 5.

As described in Chapter 5, encompassing regressions have been widely applied to evaluate volatility forecasts. They allow the researcher to assess whether forecasts are biased and/or efficient with respect to alternative forecasts. Further, the R^2 of encompassing regressions is used to rank different forecasting models. Despite its common use, this approach has certain disadvantages. Hansen (2005) criticises the use of R^2 to produce misleading forecast rankings, as the coefficient of determination ignores possible forecasting bias.[108] Another disadvantage of encompassing regressions concerns the way in which volatility forecasts are compared. Becker et al. (2007) argue that the use of encompassing regressions is equivalent to an iterative comparison of individual forecasts (e.g., BS ATM implied volatility) to alternative individual forecasts from a set of multiple forecasts (e.g., GARCH model-based forecasts from the set of time series models). Because only individual forecast comparisons are considered, the approach neglects the comparison of the individual forecasts with the complete set of alternative forecasts. Thus, the shortcomings of individual forecasts (e.g., time series models) could explain the improved forecasting performance of implied volatilities documented by several studies using encompassing regressions.[109] For this reason, encompassing regression are not used in this

[106] See Clark and McCracken (2005), p. 28

[107] The evaluation procedure applied in this study also considers certain aspects of in-sample model performance, as the model identification and selection are based on the full sample. Further, certain important, comparable studies use out-of-sample performance tests. See, e.g., Claessen and Mittnik (2002), Martens and Zein (2004), and Koopman et al. (2005).

[108] See Hansen and Lunde (2005), p. 877.

[109] See Becker et al. (2007), p. 2536.

study for the evaluation of DAX volatility forecasts.[110] Thus, the application of statistical error measures, utility-based criteria, and profit-based measures to assess the performance of volatility forecasts is discussed below.

While utility-based criteria are theoretically appealing, their practical implementation is complicated, since utility functions are typically unknown. Further, if profit-based measures instead are used to evaluate forecasting performance from an economic perspective, they require the definition of trading strategies, which imposes further assumptions regarding transaction costs, such as execution costs and bid-ask spreads.[111] Moreover, Poon and Granger (2003) argue that comparing forecast accuracy based on option pricing errors privileges forecasts derived from implied volatilities relative to predictions from time series models. The reason is that volatility prediction errors induced by misleading option pricing models are cancelled out when implied volatility is reintroduced into the option pricing formula.[112] Thus, to avoid additional assumptions when specifying utility functions or trading strategies, statistical loss functions are used to evaluate DAX volatility forecasts.[113]

Statistical Loss Functions

The most common statistical loss functions or evaluation criteria are mean error (ME), MAE, MSE, root mean square error (RMSE), and MAPE.[114] These criteria are used to rank forecasting models and select the model that reports the lowest error measure. The listed loss functions are symmetric, as they penalise positive and

[110] Becker and Clements (2008) remark on an additional disadvantage of encompassing regressions. They refer to the findings of Patton (2011), who shows that the results of encompassing regressions depend on the assumed distribution of the volatility proxy. See Becker and Clements (2008), p. 127.
[111] For instance, Chan et al. (2009) apply realised volatilities to forecast implied volatilities and investigate the economic value of their forecasting method using a dynamic trading strategy based on straddles.
[112] See Poon and Granger (2003), p. 491.
[113] Although this work considers volatility forecasts that can be used in risk management applications, specific risk management criteria (e.g., the proportion of failures) are not employed. This is because such criteria require density or quantile forecasts, which are beyond the scope of this study.
[114] See Poon and Granger (2003), p. 490.

negative errors of the same magnitude equally. To allow for asymmetric penalties, additional loss functions are proposed in the literature. For instance, Bollerslev et al. (1994) suggest the so-called QLIKE measure, which is given by

$$\text{QLIKE} = \frac{1}{H}\sum_{t=1}^{H} \log(f_t^m) + \frac{RV_t^{1/2}}{f_t^m} \tag{6.23}$$

where f_t^m denotes an individual volatility forecast series with length H from model m at time t and $RV_t^{1/2}$ represents the realised volatility at t.[115] This loss function is motivated by its implicit use for estimating volatility models based on the quasi-maximum likelihood function (e.g., GARCH models).[116] Following Harvey (1997), Kennedy (2003), and others, the choice of the evaluation criteria should be based on the ability of the measures to proxy for the economic loss resulting from the use of the forecasts.[117] Thus, the evaluation criteria selected for this study should agree with the application of DAX volatility forecasts for asset and risk management purposes.

According to West (1996), the MSE which is defined as

$$\text{MSE} = \frac{1}{H}\sum_{t=1}^{H} (f_t^m - RV_t^{1/2})^2 \tag{6.24}$$

is the most commonly used loss function.[118] It penalises extreme incorrect predictions more heavily than, e.g., the MAE, and thus is useful for applications in which extreme forecast errors are unduly more serious than small errors.[119] As extremely incorrect volatility forecast errors can lead to massive over- or under-capitalisation of financial institutions, the choice of the MSE seems meaningful for risk management applications. Further, as the MSE has been applied in many studies, its usage

[115] For instance, the QLIKE measure is applied by Becker and Clements (2008), Hansen and Lunde (2005), and Louzis et al. (2012), among others.

[116] Alternatively, Granger (1999) proposes the *LINEX* loss function initially introduced by Varian (1975).

[117] See Harvey (1997), p. 4 and Kennedy (2003), p. 362.

[118] See West (2006), p. 101.

[119] See Brooks and Persand (2003), p. 5.

allows the results of this study to be compared with the previous findings in the literature. However, in situations in which a different weighting of positive and negative forecasting errors is required, the QLIKE measure's ability to allow for asymmetric penalties renders it more appropriate.

As it is advisable to consider more than one evaluation criterion, two loss functions are considered in this study. In addition to their previously mentioned suitable characteristics, the MSE and the QLIKE are employed to evaluate DAX volatility forecasts, as Patton (2011) demonstrates that both loss functions are robust to noise in the volatility proxy. As noted above, the use of imperfect volatility proxies can lead to incorrect forecast rankings. To avoid such distortions, Patton (2011) suggests a new parametric family of loss functions that nests the MSE and the QLIKE. Having selected the evaluation criteria, the next Section presents predictability tests that examine whether two or more competing forecasts differ significantly.

Predictive Ability Tests

The commonly employed tests for evaluating the performance of two alternative prediction models are based on ratios of or differences in the above-presented statistical error measures. As the null hypothesis supposes that the predictive ability of the two models is equal, these tests are called equal predictive ability (EPA) tests. Given two squared forecast error series e_{1t}^2 and e_{2t}^2 with length H, the null hypothesis can be tested by the following F-statistic

$$F = \frac{\sum_{t=1}^{H} e_{1t}^2}{\sum_{t=1}^{H} e_{2t}^2}\,.\text{[120]} \tag{6.25}$$

Under the null hypothesis the test statistic is F-distributed with (H, H) degrees of freedom. Several assumptions must be met to employ this test. First, the test requires that the forecast errors have zero mean and are normally distributed. Further, no serial correlation in the forecast errors is allowed, and the contemporaneous

[120]The numerator of this test statistic contains the larger of the two MSE.

correlation of the two forecast error series must be zero. Because these assumptions are often violated in practice, the test statistic is not generally F-distributed.[121] Therefore, alternative test statistics have been developed that (partially) account for these violations.[122]

Diebold and Mariano (1995) and West (1996) suggest EPA tests that are valid under more general assumptions. Similar to the above F-test, Diebold and Mariano (1995) propose an EPA test that comparest two competing forecasting models. The Diebold and Mariano (DM) (1995) test is based on the sample mean of the observed loss differential series, which is given by

$$\overline{d}_{12} = \frac{1}{H}\sum_{t=1}^{H}[g(e_{1t}) - g(e_{2t})] \tag{6.26}$$

where $g(\cdot)$ represents various loss functions (e.g., MSE or asymmetric loss functions).[123] The test statistic of the DM test is given by

$$S = \frac{\overline{d}_{12}}{\sqrt{\frac{2\pi \widehat{f}_{d_{12}}(0)}{H}}} \tag{6.27}$$

where $\widehat{f}_{d_{12}}(\cdot)$ is a consistent estimate of the spectral density for the series $d_{12,t}$.[124]

As Harvey et al. (1997) demonstrate that the DM test is oversized in small samples, they develop a modified DM test statistic that performs better in small samples. The modified DM test statistic proposed by Harvey et al. (1997) is

$$S^* = \left[\frac{H + 1 - 2\tau + H^{-1}\tau(\tau - 1)}{H}\right]^{-1/2} \frac{\overline{d}_{12}}{\left(\mathrm{Var}\left(\overline{d}_{12}\right)\right)^{1/2}} \tag{6.28}$$

---

[121] For instance, multi-step-ahead forecasts can induce serial correlation in the forecast errors. Additionally, the forecast errors from two competing forecast models will typically be correlated.

[122] See Enders (2004), p. 84.

[123] The index numbers of the loss differential series $d_{12,t}$ refer to the corresponding forecasting models $m = 1, 2$.

[124] See Diebold and Mariano (1995), p. 254.

where τ denotes the forecast horizon. The variance in the serially correlated loss differentials is estimated by

$$\widehat{\text{Var}}\left(\overline{d}_{12}\right) = (\gamma_0 + 2\gamma_1 + ... + 2\gamma_q)/(H-1) \tag{6.29}$$

where γ_j is the j-th autocovariance of $\overline{d}_{12}$. Under the null hypothesis the test statistic S^* is t-distributed with $(H-1)$ degrees of freedom, provided that fairly weak conditions are satisfied.[125] In contrast to the above EPA test based on the F-statistic, the DM test and its modified version are robust to non-Gaussian and nonzero mean forecast errors. Moreover, the test does not require serially and contemporaneously uncorrelated forecast errors. However, the application of the DM tests is restricted to non-nested forecasting models.[126,127] Further, the DM test applies to predictions that do not rely on parameter estimates. As economic predictions are typically based on such estimates, West (1996) develops a framework that allows for forecast uncertainty due to parameter estimation errors.[128]

West (1996) demonstrates that the variance-covariance matrix of the loss differentials becomes more complex if parameter estimation errors are taken into account.[129] Similar to the DM test, West's (1996) framework can only be applied to non-nested models and for long series of predictions and realisations.[130] Based on formal asymptotic theory, he considers certain conditions under which parameter estimation error is asymptotically irrelevant.[131] However, the asymptotic irrelevance condition is not met in certain important cases. In particular, if rolling or fixed schemes are used and/or realisations and forecasts are correlated, the effect of parameter uncertainty

[125] See Harvey et al. (1997), pp. 281-282.
[126] See Diebold and Mariano (1995), p. 136.
[127] For the pairwise comparison of nested models, see Clark and McCracken (2001).
[128] See West (1996), p. 1067.
[129] See Fan (2010), p. 63.
[130] See West (1996), p. 1067.
[131] For instance, if the prediction period is small relative to the sample size used in the parameter estimation, asymptotic irrelevance holds. Under these conditions, inference can proceed by performing the tests suggested by Diebold and Mariano (1995).

does not vanish asymptotically and the complex variance-covariance matrix has to be computed.[132]

While the above-described test procedures consider the relative predictive accuracy of two alternative models, White (2000) generalises the Diebold and Mariano (1995) and West (1996) tests to the joint comparison of a set of multiple forecasting models against a given benchmark model.[133] White's (2000) approach is based on the composite null hypothesis that no competing model provides better forecasting results than the benchmark model. This test is called the *reality check test*, as White (2000) explicitly accounts for the effects of data-snooping. Data-snooping or data-mining can appear if a given data set is used more than once to estimate and evaluate different forecasting models, as it is possible to obtain satisfactory results due to chance.[134] Thus, if multiple forecasting models are compared, it is important to control for data-snooping biases by accounting for the correlation across the various models.[135,136]

Hansen (2005) identifies a shortcoming of the reality check test. He demonstrates that White's (2000) test is sensitive to the inclusion of poor and irrelevant forecasting models in the set of alternative prediction models. Therefore, he suggests a more powerful formulation of the test, termed the SPA test, which is based on a studentised test statistic and a sample-dependent null distribution. Similar to the reality check test, it allows for a simultaneous comparison of multiple forecasts and thus controls for data-snooping effects.[137] However, in addition to these appealing features, the SPA test has certain disadvantages. As the SPA test requires the specification of a benchmark model, it cannot be applied in cases in which no natural benchmark exists. Moreover, the rejection of the null hypothesis of the SPA test implies that one or more models outperform the benchmark. However, the test provides

[132] See West (2006), p. 111.
[133] See Corradi and Swanson (2007), p. 69.
[134] See Mariano (2004), p. 294.
[135] See Corradi and Swanson (2007), p. 69.
[136] For example, the White (2000) reality check test is employed by Awartani and Corradi (2005), who investigate the predictive ability of standard GARCH and asymmetric GARCH models.
[137] See Hansen and Lunde (2005), p. 87.

little information regarding which model is superior to the benchmark. Therefore, Hansen et al. (2003) and Hansen et al. (2011b) suggest the MCS approach, which is explained in the next Section.

The Model Confidence Set (MCS) Approach

Generally, the MCS approach can be used to examine whether several forecasts are significantly different from one an other. Specifically, the objective is to select a group of models from the initial model set that comprises the "best" forecasting models. The criterion for determining the "best" models is user-specified, e.g., the MSE. The approach is not limited to pairwise model comparisons and does not require the specification of a benchmark model.[138] Further, the MCS procedure allows the researcher to compare econometric models and more general alternatives, such as trading rules, which are not necessarily based on a specific data model.[139,140] Moreover, according to Hansen et al. (2011b), the tests results are informative regarding the informational content of the data. If the approach delivers a large set of models with equal predictive ability, this indicates that the data have limited informational content.[141] Due to these appealing features, this study employs the MCS procedure to examine whether DAX volatility forecasts based on implied volatilities are superior to volatility predictions from GARCH, ARFIMA, and HAR models. To my knowledge, this work presents the first application of the MCS approach to compare DAX volatility predictions for different forecast horizons based on realised volatility.[142] Therefore, this Section describes the approach in detail.

[138] See Hansen et al. (2003), pp. 839-843.

[139] See Hansen et al. (2011b), p. 454.

[140] Thus, the MCS approach can also be applied to implied volatilities.

[141] See Hansen et al. (2011b), pp. 459-460.

[142] In the literature, the MCS approach is applied by, inter alia, Becker and Clements (2008), Martens et al. (2009), Audrino and Hu (2011), and McAleer et al. (2013) to S&P 500 volatility, Patton and Sheppard (2009) to individual stock volatility, and Dunis et al. (2013) to the implied volatility of the EUR/USD exchange rate. In addition, Caporin and McAleer (2012) focus on model comparison and selection of univariate volatility models for financial time series, and present an empirical application of the MCS approach to one-day-ahead DAX volatility forecasts.

The objective of the MCS approach is to identify a set of models, called the MCS, which contains the "best" forecasting model for a given confidence level $(1-\alpha)$. The MCS is determined by an iterative procedure consisting of a sequence of EPA tests. At the conclusion of the sequence, the approach provides a final set $\widehat{\mathcal{M}}^*_\alpha$ of "optimal" forecasting models for a given confidence level. The final set can contain one or more models of equal predictive ability.[143]

Given an initial set of forecasting models $\mathcal{M}_0 = \{1, ..., m\}$, the MCS procedure evaluates the models according to their expected loss differentials over the sample $t = 1, ..., H$. The loss differential between models i and j in $\mathcal{M}_0$ is defined as

$$d_{ij,t} \equiv L(RV_t, f_t^i) - L(RV_t, f_t^j) \quad i, j = 1, ...m \quad t = 1, ..., H \tag{6.30}$$

and

$$d_{i.,t} \equiv L(RV_t, f_t^i) - \frac{1}{m}\sum_{j=1}^{m} L(RV_t, f_t^j) \quad i, j = 1, ...m \quad t = 1, ..., H \tag{6.31}$$

where $L(\cdot)$ represents the selected loss function (here: MSE or QLIKE). While the first sample loss statistic $d_{ij,t}$ measures the relative performance between the ith and the jth model, the second sample loss statistic $d_{i.,t}$ denotes the performance of the ith model relative to the average across the models in $\mathcal{M}$.[144]

The MCS procedure is initialised by testing the EPA hypotheses $H_0 : E(d_{ij,t}) = 0$ for the complete set of candidate models in $\mathcal{M}_0$. If the EPA hypothesis is rejected, an elimination rule is used to remove the worst-performing model from $\mathcal{M}_0$. This procedure is repeated for the remaining set of surviving models until the EPA hypothesis is accepted. The final set of surviving models provides the MCS, which is denoted by $\widehat{\mathcal{M}}^*_\alpha$.[145]

[143] See Hansen et al. (2011b), p. 453.
[144] See ibid., p. 465.
[145] See Hansen et al. (2003), pp. 843-845.

At each step of the MCS procedure, the EPA hypothesis is tested based on the test statistic

$$T_R = \max_{i,j \in \mathcal{M}} |t_{ij}| = \max_{i,j \in \mathcal{M}} \frac{|\overline{d}_{ij}|}{\sqrt{\widehat{\mathrm{Var}}(\overline{d}_{ij})}} \tag{6.32}$$

where $\overline{d}_{ij} = \frac{1}{H} \sum_{t=1}^{H} d_{ij,t}$ measures the relative performance between models i and j. If the null hypothesis is rejected, at least one model in $\mathcal{M}$ is outperformed by the other models. The worst performing model $\mathcal{M}_i$ is determined by

$$i = \arg \max_{i \in \mathcal{M}} \frac{\overline{d}_{i.}}{\sqrt{\widehat{\mathrm{Var}}(\overline{d}_{i.})}} \tag{6.33}$$

where $\overline{d}_{i.} = \frac{1}{m-1} \sum_{j \in \mathcal{M}} \overline{d}_{ij}$ represents the average performance of model i relative to the average across the models in $\mathcal{M}$. The variance estimates $\widehat{\mathrm{Var}}(\overline{d}_{ij})$ and $\widehat{\mathrm{Var}}(\overline{d}_{i.})$ are calculated using a bootstrap procedure, which is explained below. After the worst-performing model is removed from $\mathcal{M}$, the EPA test is repeated for the reduced set of models until the EPA hypothesis is accepted.[146] To obtain p-values at each stage for T_R, the implementation of the bootstrap method is described below.[147]

The distribution of T_R is approximated by bootstrapping, as the asymptotic distribution of the test statistic T_R is non-standard.[148] The bootstrap method is used to generate B bootstrap resamples for all combinations of i and j for $d_{ij,t}$. Specifically, the block bootstrap procedure that considers sequential blocks of $d_{ij,t}$ is applied to capture temporal dependencies in $d_{ij,t}$.[149] Thus, before the bootstrap procedure can be initialised, the time series dependencies in $d_{ij,t}$ must be analysed. Hansen et al. (2003) suggest fitting an autoregressive process to $d_{ij,t}$ and setting the block length

[146] See Hansen et al. (2011b), pp. 465-466.
[147] Similar results hold for $T_{\max}$.
[148] The distribution of T_R depends on the covariance structure of the forecasts. See Becker and Clements (2008), p. 128.
[149] Alternatively, the stationary bootstrap method developed by Politis and Romano (1994) can be used. In contrast to the block bootstrap procedure, this procedure is based on random block lengths. While the stationary bootstrap method ensures the stationarity of the bootstrap resamples, the variance of the statistics increases (see Becker and Clements (2008), p. 128). Thus, similar to Hansen et al. (2003), the block bootstrap method is employed in this study.

to the largest lag length.[150] Given the selected block length the bootstrap procedure operates as follows:[151]

1. Generate B block bootstraps for all forecast combinations, $d_{ij,t}^b$ for $b = 1, ..., B$.

2. Estimate the variances

$$\widehat{\text{Var}}(\overline{d}_{ij}) = \frac{1}{B}\sum_{b=1}^{B}\left(\overline{d}_{ij}^b - \overline{d}_{ij}\right)^2 \tag{6.34}$$

   and

$$\widehat{\text{Var}}(\overline{d}_{i.}) = \frac{1}{B}\sum_{b=1}^{B}\left(\overline{d}_{i.}^b - \overline{d}_{i.}\right)^2 \tag{6.35}$$

   where $\overline{d}_{ij}^b$ and $\overline{d}_{i.}^b$ denote the bootstrap counterparts of $\overline{d}_{ij}$ and $\overline{d}_{i.}$.

3. Calculate the bootstrap distribution of T_R under the null hypothesis using

$$T_R^b = \max_{i,j\in\mathcal{M}}|t_{ij}| = \max_{i,j\in\mathcal{M}}\frac{|\overline{d}_{ij}^b - \overline{d}_{ij}|}{\sqrt{\widehat{\text{Var}}(\overline{d}_{ij})}}. \tag{6.36}$$

4. Compute the p-values of the EPA test using

$$\widehat{p} = \frac{1}{B}\sum_{b=1}^{B} 1(T_R^b > T_R) \tag{6.37}$$

   with

$$1(A) = \begin{cases} 1 & \text{if } A \text{ is true} \\ 0 & \text{if } A \text{ is false.} \end{cases} \tag{6.38}$$

Equation (6.37) demonstrates that the p-value of the EPA test is calculated as the proportion of instances in which the bootstrapped test statistics are larger than the observed value of T_R. These p-values are used to determine model-specific p-values.

[150] The block bootstrap method requires that the series $d_{ij,t}$ be stationary and exhibit geometrically strong mixing. See Hansen et al. (2003), p. 846.

[151] See ibid., pp. 860-861.

To determine of model-specific p-values, it is necessary to select a significance level α. If the p-value of the EPA test exceeds α, the worst-performing model is removed from $\mathcal{M}$. In this case, the p-value of the model is equivalent to the maximum p-value of all EPA tests up to this iteration. More formally, the p-value of model i is defined by

$$\widehat{p}_i = \max_{k \leq k(i)} p(k) \tag{6.39}$$

where $p(k)$ is the p-value of the kth EPA test and $k(i)$ denotes the iteration at which model i is removed from $\mathcal{M}$. Thus, based on this definition, the first model removed from $\mathcal{M}$, receives the smallest p-value. By convention, the p-value of the surviving models is 1.[152]

Despite its appealing properties, some issues must be accounted for when the MCS approach is applied in the presence of parameter uncertainty. Hansen et al. (2011b) note that the results of the MCS procedure can also be affected by parameter estimation errors. To reduce these effects, they recommend estimating the parameters based on a rolling window. Specifically, if nested models that rely on estimated parameters are compared, they note that a careful application of the MCS approach is advisable.[153] As this study estimates various time series models to predict DAX volatility, the effects of parameter estimation errors on the MCS approach are minimised by the implementation of a rolling estimation scheme. Although Hansen et al. (2011b) report that this approach has good small sample properties, the final set $\widehat{\mathcal{M}}^*_\alpha$ may include several inferior forecasting models in finite samples. This can be explained by the MCS procedure, which only eliminates a model from the MCS when its performance is significantly below that of a competing model.[154] These aspects should be recalled during the following evaluation of DAX volatility forecasts. The next Section presents the methodology applied to examine the out-of-sample performance of implied volatilities and time series models.

[152] See Hansen et al. (2011b), pp. 462-463.
[153] In addition, they suggest certain modifications, e.g., the use of a proper test size. See ibid., p. 476.
[154] See ibid., p. 493.

6.6.3. Forecasting Methodology

The choice of the forecast horizon is related to the objective of the study. As mentioned in the introduction, the objective of this work is to provide information to investment and risk managers regarding which forecasting method delivers superior DAX volatility forecasts. Typically, these practitioners require volatility forecasts of between one day and one month. Therefore, the selected forecast horizons are one day, two weeks (or ten trading days), and one month (22 trading days).[155] First, this Section describes the construction of DAX volatility forecasts based on implied volatilities.

Generally, a rolling estimation scheme is used to generate out-of-sample forecasts.[156] To obtain a series of monthly, non-overlapping forecasts, the implied volatilities are estimated based on DAX options traded on the Wednesday immediately following the expiration date.[157] In particular, the Nadaraya-Watson estimator is employed to produce smoothed implied DAX volatilities with a remaining lifetime of precisely one month.[158] These smoothed implied volatilities that begin on December 23rd, 2003 are used as one-month-ahead DAX volatility forecasts. The one-day-ahead (two-week-ahead) forecasts based on implied DAX volatilities are calculated from DAX option prices recorded every (second) Wednesday beginning from December 23rd, 2003.[159] Next, the computation of rolling DAX volatility forecasts based on

[155] As this study does not focus on forecasting intraday volatility, I refer interested readers, for example, to Engle and Sokalska (2012).

[156] The following methodology is similar to that in Jiang and Tian (2005) and Muzzioli (2010).

[157] As the week after the expiration date is one of the most active, DAX options from the middle of this week are selected. If the Wednesday is not a trading day, the next trading day is considered. See also Muzzioli (2010), p. 567.

[158] See Section 3.1.3 for the calculation of BS ATM and model-free implied volatilities using the Nadaraya-Watson estimator.

[159] While the smoothing procedure technically allows the researcher to compute implied volatilities for options with a remaining lifetime of one or ten trading days, the same implied DAX volatility for one-month-ahead forecasts is used to predict DAX volatility for shorter forecast horizons. The reason is that DAX options with fewer than five days to maturity are excluded from the sample due to liquidity concerns. As this elimination rule also reduces the data available to construct the two-week-ahead forecasts, the procedure typically suggested in the literature is to use one-month-ahead forecasts for shorter forecast horizons. See, e.g., Martin et al. (2009), pp. 89-92.

the above-presented time series models (GARCH, ARFIMA, and HAR models) is described.

The time series models are estimated for each forecast horizon based on a rolling window of 500 trading days. To match the forecast horizon of the time series models with that of implied DAX volatilities, the estimation period for one-month-ahead forecasts ends on December 23rd, 2003. The time series models are estimated based on this sample and, then, used to calculate DAX volatility forecasts for the following month.[160] In total, the application of this estimation and forecasting scheme produces 72 one-month-ahead DAX volatility forecasts. The rolling scheme must be adjusted to calculate the one-day- and ten-day-ahead forecasts.

Similar to the procedure for one-month-ahead forecasts, the estimation periods for shorter forecast horizons end on December 23rd, 2003. The predicted DAX volatilities refer to the next one or ten trading days. In the following, the estimation period for the time series model is shifted by one (ten days) towards the end of the sample and corresponding DAX volatility forecasts are computed. In summary, this procedure yields 156 two-week and 1535 one-day-ahead DAX volatility forecasts.[161] Finally, the realised volatilities used to measure the accuracy of the volatility forecasts are calculated for each forecast horizon.[162]

6.7. Evaluation of the Forecasting Results

This Section presents the evaluation of the DAX volatility forecasting results of the employed time series models and implied volatilities. In addition to the individual forecasts, this study also considers combined forecasts because forecast combinations

[160] The second sample begins on January 30th, 2002 and ends on January 21st, 2004, and so on.

[161] The following Sections consider the DAX volatility forecasts that are generated based on the GARCH(1,1) model because the GARCH(1,2) model produces considerably higher out-of-sample forecast errors.

[162] Specifically, realised volatility is computed as the average realised volatility over the forecast horizon. See also Becker and Clements (2008), p. 126.

have been found to outperform individual forecasting models in many areas.[163] For instance, Becker and Clements (2008) report that a combination of model based forecasts of S&P 500 volatility provides better prediction results than a wide range of individual forecasts, including implied volatility. The literature offers several explanations for the empirical success of combined forecasts. First, the combination of individual forecasts is an attractive strategy due to diversification gains. Second, if structural breaks are difficult to detect in the data generating process, combined forecasts based on models that adapt to these changes differently provide better forecasts on average.[164] Third, relative to individual forecasting models, a combination of forecasts can reduce misspecification biases and measurement errors. Despite these advantages, several arguments against combining forecasts are also mentioned in the literature.[165]

While non-stationarities motivate the use of combined forecasts (recall the second argument above), they can also induce stability problems in the weights used in the combinations. Further, if small samples are used to calculate numerous forecasts, estimation errors can complicate the determination of the combination weights.[166] To avoid these problems and maintain the above-mentioned number of volatility forecasts for evaluation process,[167] equal weights are used to combine individual DAX volatility forecasts. Using equal weights provides a natural benchmark for combining forecasts. Furthermore, empirical studies demonstrate that equal weighting is unlikely to be outperformed by other weighting schemes.[168,169] Therefore, this study uses equal weights to combine individual DAX volatility forecasts based on time series models and/or implied volatilities. Finally, the random walk model completes the set of examined forecasting approaches, as it is typically used as a benchmark.

[163] See, e.g., the comprehensive studies by Makridakis and Hibon (2000), Stock and Watson (1999), and Marcellino (2004).
[164] For an explanation see Pesaran and Timmermann (2007).
[165] See Timmermann (2006), pp. 137-138.
[166] See ibid., p. 139.
[167] Specifically, the one-month ahead DAX volatility forecast series consists of 72 data points.
[168] See Timmermann (2006), p. 193.
[169] To understand these empirical findings, Timmermann (2006) provides general conditions under which equal weights are optimal in a population sense. See ibid., pp. 148-150.

Because the sample period contains the 2008 financial crisis, it is possible to analyse the impact of the crisis on model forecasting performance, which is of particular interest for academics and practitioners. The literature suggests various approaches to cope round with extremely volatile sample periods. For instance, Fleming (1998) excludes all data surrounding the October 1987 stock market crash, as model parameter estimates are highly influenced by this event.[170] Alternatively, using a GARCH framework, Blair et al. (2001) introduce dummy variables in the mean and variance equations to capture the effects of the October 1987 crash. Further, Christensen and Prabhala (1998), who observe a regime shift around the 1987 crash, perform a pre-crash and a post-crash subperiod analysis.[171] Due to the forecasting methodology employed and the occurrence of the financial crisis during the last third of the sample period, the forecasting performance of the implied volatilities and time series models is evaluated for the full sample (from 2002 to 2009) and the subperiod excluding the two most volatile months of the financial crisis (September and October 2008) separately.[172] First, the following Section discusses the evaluation results for the one-day-ahead DAX volatility forecasts.

6.7.1. One-day-ahead Forecasts

The use of the MCS approach requires that all loss differentials d_{ij} are stationary. Therefore, the ADF test is performed to test for non-stationarity in the loss differentials. Table B.5 in the Appendix presents the results of the ADF test, the null hypothesis of which states that each series follows a unit-root process.[173,174] As the null hypothesis can be rejected for all loss differentials at the 1% level, the MCS method can be implemented to evaluate the performance of the DAX volatility forecasts.

[170] See Fleming (1998), p. 323.
[171] See Christensen and Prabhala (1998), pp. 141-147.
[172] In the following, this sample is called “subperiod ex financial crisis 2008”.
[173] The alternative hypothesis is that the series are generated by a stationary process without drift.
[174] Using the results of the dependency analysis below, the ADF test includes up to 15 lags.

As described in the previous Section, the MCS approach uses the block bootstrap method to generate the distributions of the test statistics. Similar to Hansen et al. (2003), the block length is determined based on the lag length of the autoregressive processes fitted to the loss differentials. Specifically, the block length is set equal to the largest lag length of the estimated autoregressive processes. Table B.8 in the Appendix provides the estimation results for the autoregressive processes fitted to the daily loss differentials.[175] The table indicates that the maximum lag length is 15. Thus, the block length is set to 15. To ensure that the results are not influenced by the actual draws, 10,000 bootstrap resamples are generated.[176] The MCS approach is employed based on these parameters, and the results regarding the MSE criteria are presented in Table 6.10.

When the full sample is considered, the combined forecast of BS ATM implied volatility and the HAR model provides the lowest MSE. In addition to this combination, the combined forecast of BS ATM implied volatility and the ARFIMA(1,d,1) model and the individual forecasts from the HAR and the ARFIMA(1,d,1) exhibit low MSE values. In contrast, the highest MSE values are observed for the individual forecasts based on the implied volatilities and the GARCH models. While the relatively poor performance of the volatility forecasts based on the implied volatilities can be explained by the mismatch between the forecast horizon and the maturity, the inferior performance of the GARCH models is surprising. Typically, GARCH models provide good prediction results over short-term periods. In the following, the MCS approach is applied to examine whether the two combined forecasts and the individual forecasts from the HAR and ARFIMA(1,d,1) models significantly outperform forecasts based on the implied volatilities and GARCH models.

Table 6.10 indicates that the MCS comprises four combined forecasts and the individual forecasts from the HAR and ARFIMA(1,d,1) models. Therefore, each of the

[175] The lag length of the autoregressive processes is determined based on the SIC. As outliers distort the sample autocorrelation structure of the daily loss differentials, observations with realised volatilities above 75% are excluded from the data set (approximately 1% of the observations). For the effects of outliers on sample autocorrelations, see Chan (1995).

[176] The same number of resamples is used for the two-weeks-ahead and one-month-ahead forecasts.

Table 6.10.: MCS results for one-day-ahead forecasts (loss function: MSE)

model	MSE	rank (MSE)	MCS p-value
Panel A: full sample			
garch11	0.517%	(10)	0.3%
egarch21	0.521%	(11)	1.3%
arfima11	0.403%	(3)	93.7%*
har	0.400%	(2)	93.7%*
mfv	0.594%	(12)	0.3%
bsatm	0.487%	(9)	1.3%
garch11+arfima11	0.417%	(5)	37.8%*
garch11+har	0.417%	(5)	81.8%*
bsatm+garch11	0.462%	(7)	0.3%
bsatm+arfima11	0.403%	(3)	92.9%*
bsatm+har	0.399%	(1)	100.0%*
rw	0.473%	(8)	1.3%
Panel B: subperiod ex financial crisis 2008			
model	MSE	rank (MSE)	MCS p-value
garch11	0.349%	(11)	0.3%
egarch21	0.310%	(8)	0.5%
arfima11	0.256%	(3)	72.8%*
har	0.256%	(3)	72.8%*
mfv	0.397%	(12)	0.3%
bsatm	0.330%	(10)	0.3%
garch11+arfima11	0.268%	(5)	30.3%*
garch11+har	0.268%	(5)	72.8%*
bsatm+garch11	0.307%	(7)	0.3%
bsatm+arfima11	0.255%	(2)	72.8%*
bsatm+har	0.252%	(1)	100.0%*
rw	0.313%	(9)	0.5%

Source: own calculations.

Note: The forecasts in $\widehat{\mathcal{M}}^*_{90\%}$ are identified by one asterisk.

four combined forecasts contains one of the individual forecasts from the HAR and ARFIMA(1,d,1) models. As the individual forecasts based on the implied volatilities and the GARCH models do not belong to the MCS at the 5% level, it follows that they provide significantly inferior prediction results relative to the MCS models. The next paragraph presents the prediction results for the subsample that excludes the two most volatile months of the 2008 financial crisis.

As expected, all examined forecasting methods exhibit considerably lower MSE values when the reduced subsample is considered. With respect to the ranking of the volatility forecasts, the findings for the reduced subsample are similar to those for the full sample (see Table 6.10). As in the case of the full sample, the two combined forecasts of BS ATM implied volatility and the HAR or the ARFIMA(1,d,1) model, followed by the individual forecasts of the HAR and ARFIMA(1,d,1) models produce the lowest MSE values. Furthermore, the forecasts based on the implied volatilities and GARCH(1,1) exhibit the highest MSE errors. Despite these similarities, the evaluation results of these two samples are not identical. Interestingly, the relative ranking of the GARCH(1,1) and EGARCH(2,1) models changes when the two most volatile months of the crisis are eliminated. This implies that the higher MSE documented for the EGARCH(2,1) model (relative to the GARCH(1,1) model) for the full sample is due to the extreme market movements in autumn 2008. This finding agrees with the results of Bluhm and Yu (2001), who report that asymmetric GARCH models (in particular the GJR-GARCH and the EGARCH model) provide better one-day-ahead DAX volatility forecasts than the standard GARCH model.[177]

The MCS results for the reduced subsample are presented in Table 6.10 and indicate that the composition of the MCS remains constant. Thus, except for minor changes in individual rankings, the 2008 financial crisis affects the one-day-ahead predictive abilities of all forecasting methods considered, as measured by the MSE, in the same

[177]See Bluhm and Yu (2001), pp. 13-15.

direction. After discussing model forecasting performance based on the MSE, the results for the asymmetric loss function, QLIKE, are outlined below.

Table 6.11.: MCS results for one-day-ahead forecasts (loss function: QLIKE)

Panel A: full sample			
model	QLIKE	rank (QLIKE)	MCS p-value
garch11	4.591%	(10)	0.0%
egarch21	4.524%	(9)	0.0%
arfima11	3.566%	(1)	100.0%*
har	3.567%	(2)	98.7%*
mfv	6.036%	(12)	0.0%
bsatm	4.874%	(11)	0.0%
garch11+arfima11	3.716%	(6)	29.0%*
garch11+har	3.686%	(5)	38.9%*
bsatm+garch11	4.395%	(7)	0.0%
bsatm+arfima11	3.669%	(4)	38.9%*
bsatm+har	3.606%	(3)	88.4%*
rw	4.482%	(8)	0.0%
Panel B: subperiod ex financial crisis 2008			
model	QLIKE	rank (QLIKE)	MCS p-value
garch11	4.480%	(10)	0.0%
egarch21	4.297%	(7)	0.2%
arfima11	3.448%	(1)	100.0%*
har	3.468%	(2)	73.9%*
mfv	5.946%	(12)	0.0%
bsatm	4.818%	(11)	0.2%
garch11+arfima11	3.604%	(6)	27.8%*
garch11+har	3.582%	(5)	31.4%*
bsatm+garch11	4.318%	(8)	0.2%
bsatm+arfima11	3.576%	(4)	31.4%*
bsatm+har	3.520%	(3)	73.9%*
rw	4.436%	(9)	0.2%

Source: own calculations.
Note: The forecasts in $\widehat{\mathcal{M}}^*_{90\%}$ are identified by one asterisk.

Using the QLIKE criterion to rank the DAX volatility forecasts yields similar results to the MSE loss function (see Table 6.11). As the MCS contains the same forecasting methods, this confirms the findings above regarding the superior predictive ability of the individual forecasts from the ARFIMA(1,d,1) and the HAR model and those of the combined forecasts including one of these individual forecasts.[178] In all comparisons, the top position in the volatility forecast ranking changes if the forecasts are ranked according to the QLIKE criterion. While the lowest MSE is observed for the combined forecast based on BS ATM implied volatility and the HAR model, the lowest QLIKE is received for the individual forecast from the ARFIMA(1,d,1) model.[179] In accordance with the findings for the MSE, the highest values of the QLIKE measure are recorded for the individual forecasts based on the implied volatilities and the GARCH(1,1) model. When the forecasting results for the reduced subsample "subperiod ex financial crisis 2008" are considered, the MCS still contains the same models. Thus, the findings regarding the superior performance of these forecasting models also holds for the QLIKE measure and the reduced subsample.

Overall, the composition of the MCS does not change when the loss function is changed (MSE to QLIKE) or a different sample is considered (the full sample or reduced subsample). Accordingly, the combined forecasts that contain one of the individual forecasts from the HAR or ARFIMA(1,d,1) models and the both individually produce forecasts that are superior to competing model-based forecasts. This finding demonstrates that the ability of the HAR and ARFIMA(1,d,1) models to capture long memory dependencies in realised DAX volatilities seems important for predicting short-term DAX volatility. The results regarding whether combined forecasts are superior to individual forecasts are mixed. While in the case of the MSE criterion the combination of the individual forecasts from BS ATM implied volatility and the HAR or ARFIMA(1,d,1) model (slightly) reduces the forecast errors relative to the individual forecasts (HAR and ARFIMA(1,d,1)), this effect cannot be observed for the QLIKE measure.[180]

[178] Similar to the MCS results for the MSE criterion, the MCS based on the QLIKE criterion contains the combined forecasts of the GARCH(1,1) model and the HAR or the ARFIMA(1,d,1) model.

[179] The combined forecast of BS ATM implied volatility and the HAR model also produces low QLIKE errors.

[180] Note that according to the MCS approach, the combined forecast that exhibits the lowest MSE does not significantly outperform the individual forecasts from the long memory models.

These results support the findings of Lazarov (2004), who reports that the standard ARFIMA model and the ARFIMA model enhanced with implied variance provide the best DAX volatility forecasts.[181] Although this study does not consider extended ARFIMA models, the results of Lazarov (2004) are related to this work. Similar to the extended ARFIMA model suggested by Lazarov (2004), the combined forecast of BS ATM implied volatility and the HAR model suggested by this study is based on the information from using realised and implied volatilities to predict DAX volatility. Further, this study extends the findings of Lazarov (2004), as his evaluation approach does not allow for the investigation of whether the (extended) ARFIMA models provide significantly better DAX volatility forecasts than the (extended) GARCH models. Specifically, based on the MCS approach, this study demonstrates that the ARFIMA and related HAR model significantly outperform the GARCH and the EGARCH models. The following Section presents the evaluation results for the two-weeks-ahead DAX volatility forecasts.

6.7.2. Two-weeks-ahead Forecasts

To apply the MCS approach to evaluate the two-weeks-ahead forecasts, the ADF test is used to examine whether the loss differentials d_{ij} are stationary. The results of the ADF test, which are reported in Table B.6 in the Appendix, show that the null hypothesis of non-stationarity can be rejected for all loss differentials at the 1% level.[182] Thus, the MCS method can be employed to assess the predictive ability of the selected forecasting models. As in the case of the one-day-ahead forecasts, the block length of the bootstrapping method is determined based on autoregressive processes fitted to the bi-weekly loss differentials.[183] The estimation results for these autoregressive processes imply a block length of four (see Table B.10). The outcome

[181] See Lazarov (2004), pp. 60-61.

[182] Based on the results of the following temporal dependency analysis, the ADF test is performed using four lags.

[183] Similar to the procedure for one-day-ahead loss differentials, the lag length of the autoregressive processes are determined based on the SIC. Here, no outliers have to be excluded from the data set.

of the MCS approach and the two-weeks-ahead forecast ranking based on the MSE criterion are presented in Table 6.12.

Table 6.12.: MCS results for two-weeks-ahead forecasts (loss function: MSE)

Panel A: full sample			
model	MSE	rank (MSE)	MCS p-value
garch11	0.494%	(9)	8.6%
egarch21	0.624%	(12)	8.6%
arfima11	0.499%	(10)	8.6%
har	0.504%	(11)	8.6%
mfv	0.471%	(8)	8.6%
bsatm	0.394%	(1)	100.0%*
garch11+arfima11	0.439%	(5)	77.3%*
garch11+har	0.451%	(7)	8.6%
bsatm+garch11	0.401%	(4)	99.6%*
bsatm+arfima11	0.400%	(2)	99.6%*
bsatm+har	0.400%	(2)	99.6%*
rw	0.448%	(6)	96.6%*
Panel B: subperiod ex financial crisis 2008			
model	MSE	rank (MSE)	MCS p-value
garch11	0.322%	(12)	4.5%
egarch21	0.257%	(10)	3.3%
arfima11	0.206%	(4)	8.2%
har	0.205%	(3)	6.5%
mfv	0.273%	(11)	4.5%
bsatm	0.223%	(5)	6.5%
garch11+arfima11	0.225%	(6)	4.5%
garch11+har	0.229%	(7)	4.5%
bsatm+garch11	0.237%	(8)	4.5%
bsatm+arfima11	0.175%	(2)	28.1%*
bsatm+har	0.171%	(1)	100.0%*
rw	0.246%	(9)	4.5%

Source: own calculations.

Note: The forecasts in $\widehat{\mathcal{M}}^*_{90\%}$ are identified by one asterisk.

In the full sample, the individual forecast based on BS ATM implied volatility provides the lowest MSE prediction error for two-weeks-ahead DAX volatility forecasts. In contrast to the findings for one-day-ahead forecasts, the ARFIMA(1,d,1) and the HAR model produce relatively high MSE values when applied to the full sample.

While the superior performance of BS ATM implied volatility can be explained by the better match of forecast horizon and an option's time to maturity, the unexpectedly poor results of the HAR and the ARFIMA(1,d,1) model must be further analysed using the MCS results for the reduced subsample presented in Table 6.12. The findings for the subsample demonstrate that the high MSE values observed for the long memory models are due to the impact of the two most volatile months of the financial crisis.[184] Similar to the results for one-day-ahead DAX volatility forecasts, the combined forecasts that include the individual forecasts of BS ATM implied volatility and the HAR or the ARFIMA(1,d,1) model exhibit low MSE errors and, in contrast, the GARCH models produce high MSE values.

Another notable result is that the DAX volatility forecasts based on BS ATM implied volatility produce a considerably lower MSE than model-free volatility based forecasts. Additionally, the results of the MCS approach show that the individual forecast based on BS ATM implied volatility provides significantly lower MSE prediction errors than does that based on model-free volatility. Although model-free volatility, by construction, uses a larger information set than BS implied volatility, this additional information does seem not to be relevant for the prediction of two-weeks-ahead DAX volatility.

While the MCS for the full sample contains six models, including the random walk model, the MCS for the reduced sample is more compact. In particular, the MCS for the reduced sample shows that the combined forecasts of BS ATM implied volatility and the HAR, respectively, the ARFIMA(1,d,1), model provide significantly better forecasts than the other prediction methods. Further, the reduction of the MCS reflects the considerable impact of the financial crisis on forecasting two-weeks-ahead DAX volatility.

Whereas the MCS based on the MSE loss function contains multiple models and differs across the two samples, the MCS results for the QLIKE function indicate that the combined forecast of BS ATM implied volatility and the HAR model provides

[184] For the reduced subsample, the HAR and the ARFIMA(1,d,1) model show relatively low MSE values compared with the other forecasting models.

Table 6.13.: MCS results for two-weeks-ahead forecasts (loss function: QLIKE)

Panel A: full sample

model	QLIKE	rank (QLIKE)	MCS p-value
garch11	3.687%	(9)	1.4%
egarch21	4.180%	(12)	2.3%
arfima11	3.710%	(10)	1.7%
har	3.587%	(8)	1.7%
mfv	4.094%	(11)	1.7%
bsatm	3.220%	(4)	4.8%
garch11+arfima11	3.310%	(6)	3.1%
garch11+har	3.234%	(5)	3.1%
bsatm+garch11	3.123%	(3)	3.1%
bsatm+arfima11	2.792%	(2)	4.8%
bsatm+har	2.674%	(1)	100.0%*
rw	3.475%	(7)	3.1%

Panel B: subperiod ex financial crisis 2008

model	QLIKE	rank (QLIKE)	MCS p-value
garch11	3.317%	(11)	0.1%
egarch21	3.284%	(9)	0.7%
arfima11rv	3.119%	(8)	2.6%
har	3.080%	(7)	2.6%
mfv	3.847%	(12)	0.7%
bsatm	3.024%	(6)	2.6%
garch11+arfima11rv	2.866%	(4)	3.0%
garch11+har	2.811%	(3)	3.0%
bsatm+garch11	2.872%	(5)	2.6%
bsatm+arfima11rv	2.433%	(2)	4.5%
bsatm+har	2.333%	(1)	100.0%*
rw	3.292%	(10)	0.7%

Source: own calculations.

Note: The forecasts in $\widehat{\mathcal{M}}^*_{90\%}$ are identified by one asterisk.

the best DAX volatility predictions across both samples (see Table 6.13). In addition to this forecast combination, the related combined forecast that consists of BS ATM implied volatility and the ARFIMA(1,d,1) model provides low, but not significantly lower, QLIKE values. As for the MSE criterion, the forecasts based on BS ATM implied volatility show lower (QLIKE) prediction errors than model-free volatility forecasts, and the GARCH models exhibit comparably high QLIKE errors.

In the literature, Claessen and Mittnik (2002) also compare the predictive ability of time series models and implied volatility for two-weeks-ahead DAX volatility forecasts. Moreover, considering individual forecasts (e.g., the GARCH(1,1) model or an extended version of the GARCH model to which BS ATM implied volatility is added as an external regressor), they also examine the performance of combined forecasts. Their analysis shows that combining forecasts does not necessarily improve forecasting performance.[185] However, they find that a combination of BS ATM implied volatility and the GARCH(1,1) model with weekend effects provides lower prediction errors than the individual forecasts, including the extended GARCH model.[186,187]

This study partially confirms the findings of Claessen and Mittnik (2002). Based on the MSE loss function, this work shows that the combined forecast of BS ATM implied volatility and the standard GARCH(1,1) model provides lower prediction errors than the individual forecasts from the GARCH models. However, in contrast to Claessen and Mittnik (2002), this study reports a lower MSE for BS ATM implied volatility based forecasts than for the combined forecast of BS ATM implied volatility and the GARCH(1,1) model. Further, this study suggests that the combined forecasts of BS ATM implied volatility and long memory models provide superior forecasting results in (not extreme volatile) market situations.

[185] In particular, they combine a moving-average model, a random walk model, an AR(15) model, BS ATM implied volatility, and a GARCH(1,1) model.

[186] See Claessen and Mittnik (2002), pp. 314-320.

[187] This result is based on the application of the MSE criterion. In addition to the above-mentioned combined forecast, Claessen and Mittnik (2002) report that the combination of implied volatility with a constant term to correct for mean bias produces the lowest MSE. See ibid., p. 316.

A comparison of GARCH and ARFIMA models can be found in Lazarov (2004). He reports better predictive performance for (enhanced) ARFIMA models compared with GARCH or extended GARCH models. This study confirms the better performance of ARFIMA relative to GARCH models for the reduced sample.

However, while this study documents relatively high MSE values for the GARCH models, Raunig (2006) finds that an asymmetric GARCH model provides lower MSE values than the standard GARCH(1,1) model and BS ATM implied volatility.[188] This study confirms the results of Raunig (2006) with respect to the better performance of asymmetric GARCH models (here: the EGARCH model) relative to the standard GARCH(1,1) model for the reduced sample but contradicts Raunig's (2006) findings concerning the inferior performance of BS ATM implied volatility. These different results cannot be clearly attributed to the samples considered because his experiment partly overlaps with the sample period of this study. However, in contrast to this work, he uses squared daily returns and not realised volatility to measure volatility, which can explain the different outcomes. Next, the evaluation results for one-month-ahead DAX volatility forecasts are presented.

6.7.3. One-month-ahead Forecasts

The results of the ADF test presented in Table B.7 demonstrate that the non-stationary hypothesis can be rejected for all loss differentials at the 1% level. Therefore, the MCS method can be applied to evaluate the one-month-ahead DAX volatility forecasts.[189] Further, the estimation results for the autoregressive processes fitted to the monthly loss differentials show that the maximum lag length of the autoregressive processes is two (see Table B.11). Using these results to employ the MCS approach leads to the results summarised in Table 6.14 for the MSE loss function.

The comparison of the MSE rankings for one-month-ahead and two-weeks-ahead forecasts shows that some results are identical. First, the combined forecasts of

[188] See Raunig (2006), p. 371.
[189] The ADF test is performed for two lags.

Table 6.14.: MCS results for one-month-ahead forecasts (loss function: MSE)

Panel A: full sample

model	MSE	rank (MSE)	MCS p-value
garch11	0.693%	(11)	9.7%
egarch21	1.021%	(12)	11.1%*
arfima11	0.678%	(10)	11.1%*
har	0.622%	(8)	45.9%*
mfv	0.601%	(6)	45.9%*
bsatm	0.559%	(2)	92.0%*
garch11+arfima11	0.628%	(9)	45.9%*
garch11+har	0.615%	(7)	45.9%*
bsatm+garch11	0.596%	(5)	45.9%*
bsatm+arfima11	0.575%	(3)	87.0%*
bsatm+har	0.551%	(1)	100.0%*
rw	0.584%	(4)	92.0%*

Panel B: subperiod ex financial crisis 2008

model	MSE	rank (MSE)	MCS p-value
garch11	0.363%	(12)	3.9%
egarch21	0.320%	(11)	2.3%
arfima11	0.211%	(3)	53.1%*
har	0.218%	(4)	53.1%*
mfv	0.260%	(9)	53.1%*
bsatm	0.245%	(6)	41.7%*
garch11+arfima11	0.240%	(5)	53.1%*
garch11+har	0.251%	(7)	53.1%*
bsatm+garch11	0.275%	(10)	2.3%
bsatm+arfima11	0.189%	(1)	100.0%*
bsatm+har	0.192%	(2)	82.1%*
rw	0.255%	(8)	3.9%

Source: own calculations.

Note: The forecasts in $\widehat{\mathcal{M}}^*_{90\%}$ are identified by one asterisk.

BS ATM implied volatility and the HAR, or the ARFIMA(1,d,1), model and the individual forecast based on BS ATM implied volatility provide the lowest MSE values in the full sample. Second, while the individual forecasts of the HAR and the ARFIMA(1,d,1) produce high MSE prediction errors when applied to the full sample, both forecasting models provide considerably lower MSE values in the reduced sample. Third, BS ATM implied volatility provides better forecasting performance than model-free volatility in both samples. Fourth, the combined forecasts of BS ATM implied volatility and the HAR or the ARFIMA(1,d,1) model exhibit the best forecasting performance in the reduced sample. Finally, the GARCH models produce relatively high MSE errors regardless of the sample considered.

However, the MCS results for one-month-ahead and two-weeks-ahead DAX volatility forecasts exhibit some differences. Whereas the MCS for two-weeks-ahead forecasts shrinks to one, respectively, two models, the MCS for the one-month-ahead forecasts still comprises eight, respectively, ten models when the reduced sample is considered. In particular, the MCS approach demonstrates, that based on the MSE criterion, both GARCH models, the combined forecast of BS ATM implied volatility and the GARCH(1,1) model, and the random walk model, are significantly outperformed by the other forecasting models when the effects of the financial crisis are (partly) eliminated. Below, the evaluation results based on the QLIKE criterion are discussed.

Assessing the performance of the forecasting models using the QLIKE criterion yields a ranking that is similar to the previous ranking based on the MSE (see Table 6.15).[190] Although the combined forecasts of BS ATM implied volatility and the ARFIMA(1,d,1) or the HAR model deliver considerably lower QLIKE values, the MCS barely fails to verify the superior predictive ability of both forecasting methods at the 10% level. There are two potential explanations for the failure of

[190] However, although the results are closely related, there are some differences. First, the improved forecasting performance of long memory models in the reduced sample is observed in a weaker form for the QLIKE criterion. Second, BS ATM implied volatility also provides one of the best DAX volatility forecasts when the two most volatile months of the 2008 crisis are removed. The good forecasting results observed for the combined forecast of BS ATM implied volatility and the GARCH(1,1) model in both samples can be attributed to the remarkable results of BS ATM implied volatility.

Table 6.15.: MCS results for one-month-ahead forecasts (loss function: QLIKE)

Panel A: full sample

model	QLIKE	rank (QLIKE)	MCS p-value
garch11	4.756%	(9)	3.9%
egarch21	7.125%	(12)	4.3%
arfima11	5.170%	(11)	3.9%
har	4.769%	(10)	4.3%
mfv	4.408%	(7)	42.8%*
bsatm	3.932%	(2)	53.2%*
garch11+arfima11	4.538%	(8)	42.8%*
garch11+har	4.342%	(5)	42.8%*
bsatm+garch11	4.092%	(4)	42.8%*
bsatm+arfima11	3.950%	(3)	52.1%*
bsatm+har	3.714%	(1)	100.0%*
rw	4.366%	(6)	52.1%*

Panel B: subperiod ex financial crisis 2008

model	QLIKE	rank (QLIKE)	MCS p-value
garch11	3.663%	(10)	0.4%
egarch21	4.132%	(12)	1.4%
arfima11rv	3.644%	(9)	10.2%*
har	3.412%	(7)	10.2%*
mfv	3.530%	(8)	10.2%*
bsatm	3.113%	(3)	10.2%*
garch11+arfima11rv	3.261%	(6)	10.2%*
garch11+har	3.118%	(4)	10.2%*
bsatm+garch11	3.148%	(5)	10.2%*
bsatm+arfima11rv	2.820%	(2)	16.9%*
bsatm+har	2.644%	(1)	100.0%*
rw	3.717%	(11)	10.2%*

Source: own calculations.

Note: The forecasts in $\widehat{\mathcal{M}}^*_{90\%}$ are identified by one asterisk.

the MCS approach to identify one or at least a small set of superior forecasting models for both loss functions in the case of one-month-ahead forecasts. The first argument refers to the different sample sizes that are used for the evaluation of the DAX volatility forecasts. While the evaluation of one-day-ahead and two-weeks-ahead DAX volatility forecasts is based on large samples, the sample size for the one-month-ahead forecasts is much lower due to the non-overlapping construction principle employed. In addition to the sample size, the larger number of models in the MCS for one-month-ahead forecasts potentially reflects that increasing noise dominates the predictive ability of nearly all forecasting models over longer forecast horizons. In the following, the evaluation results for one-month-ahead DAX volatility forecasts are compared with the findings in the literature.

Similar to their findings for two-weeks-ahead forecasts, Claessen and Mittnik (2002) report that based on the MSE criterion the combined forecast of BS ATM implied volatility and the GARCH(1,1) model with weekend effects outperforms the individual forecasts.[191] However, this study demonstrates, that the combined forecast only produces a lower MSE than the GARCH(1,1) model.[192] In contrast, the weaker performance of BS ATM implied volatility in comparison to the GARCH(1,1) model documented by Claessen and Mittnik (2002) cannot be confirmed in this study.[193] Instead, the findings of this work are in line with those of Raunig (2006), who reports that BS ATM implied volatility exhibits lower MSE values for 30-days-ahead DAX volatility forecasts than the GARCH(1,1) and the GJR-GARCH(1,1) model.[194]

Further, the results of this study agree with a related paper published by Muzzioli (2010) who also uses non-overlapping one-month-ahead DAX volatility forecasts. According to Muzzioli (2010), both option-based volatility forecasts are better predictors of future realised volatility than the GARCH(1,1) model. In addition, she reports that BS implied volatility exhibits better performance than model-free volatil-

[191] See Claessen and Mittnik (2002), p. 317.

[192] In particular, the MCS results presented in Table 6.14 show that, in the full sample, the MSE prediction error of the combined forecast is significantly lower than for the GARCH(1,1) model.

[193] Note, that Claessen and Mittnik (2002) only consider a small sample of 26 four-week forecasts. See ibid., p. 313.

[194] See Raunig (2006), p. 371.

ity.[195,196] However, Muzzioli (2010) does not consider DAX volatility forecasts based on long memory models and combined forecasts. The above findings suggest that these forecasting approaches provide superior prediction results that should be taken into account. Hence, this thesis extends the results of Muzzioli (2010).

[195]See Muzzioli (2010), pp. 581-582.

[196]In contrast, Schöne (2010) and Tallau (2011) report that the information content of the VDAX-New is higher than that of the VDAX. Note that the VDAX is calculated from BS ATM implied volatilities and the VDAX-New is based on the concept of Demeterfi et al. (1999), which is identical with the model-free implied volatility approach developed by Britten-Jones and Neuberger (2000) (see Tallau (2011), p. 52). However, because their evaluation approach uses the Mincer-Zarnowitz regression, which is not directly comparable with the approach employed in this study, I refrain from a more detailed comparison of the results.

7. Conclusion

Although forecasting DAX volatility is crucial to option pricing, investment management, and risk management, a comprehensive overview on the performance of volatility prediction models for the German stock market does yet not exist. Furthermore, many risk models failed during the global financial crisis of 2008. Thus, a study on the forecasting ability of different DAX volatility prediction models that cover episodes of turmoil will provide important information for practitioners and academics. While the existing studies only perform isolated comparisons of forecasting models, a broad range of approaches are used in this study to predict DAX volatility.

Moreover, recently developed forecast evaluation approaches that consider data snooping effects have not been applied to the evaluation of DAX volatility forecasts. In addition, although some empirical studies provide evidence for the existence of structural breaks in financial time series, the associated effects on the prediction of DAX volatility are not examined.

The intent of this study is to close these research gaps and to provide information regarding which forecasting method delivers superior DAX volatility forecasts. Before the empirical analysis of the DAX IVS is performed to detect the dynamic regularities that can be used to predict volatility, an overview of the current research on the stylised facts of DAX implied volatilities is presented in Chapter 2.

In the literature, most empirical studies present evidence for the existence of a post-crash DAX volatility skew and a DAX volatility term structure that are similar to other options markets.[1] Fengler et al. (2002), Wallmeier (2003), and Äijö (2008) document different shapes and considerable time variation of the DAX IVS. Furthermore, Fengler (2012) finds that the DAX IVS exhibits a systematic dynamic pattern that should be considered when predicting implied volatility. Because the objective of this study is to predict DAX volatility, the structure and time variation of the DAX IVS with sufficient regard to the financial crisis of 2008 are analysed in Chapter 3. The behaviour of the DAX IVS is investigated for three different subsamples because different volatility regimes occurred during the sample period.[2]

The empirical analysis demonstrates that on average, a non-flat DAX IVS existed during the sample period from 2002 to 2009. The average DAX IVS shows that the DAX volatility smile is steepest for short-term options and flattens out with increasing maturity. The volatility of DAX implied volatilities is higher for DAX short-term options than for DAX options with longer maturities. Furthermore, the findings show that the slope of the DAX volatility smile changes during the sample period, particularly during volatile market periods, e.g., the financial crisis of 2008.[3] Similarly, the term structure varied considerably during the sample period, and the movements across the maturities are closely related.

Considering the volatility regimes, the slope of the term structure seems to be positive during market periods with normal volatility levels and negative in turbulent market phases. The change in the slope during volatile periods has not been previously documented in the literature for DAX options.[4] Moreover, the times series of DAX implied volatilities suggest that the changes of the series are highly correlated across maturity and moneyness. In particular, the series show that DAX implied volatilities tend to move upwards (downwards) during turbulent (stable) market pe-

[1] See, e.g., Wallmeier (2003), Hafner (2004), and Fengler et al. (2007).

[2] The first sample considers the turbulent market phase at the beginning of the sample period (January 2nd, 2002 to May 2nd, 2003), the second subsample comprises a long, stable upturn period (May 5th, 2003 to August 8th, 2007), and the last subsample covers the financial crisis of 2008 (August 9th, 2007 to December 30th, 2009).

[3] In particular, a volatility skew can be observed in the first and third subsamples, whereas a volatility smile occurred during the second subsample.

[4] For instance, Fengler (2004) reports a relatively flat average term structure for DAX ATM options in 2001 when the dot-com crisis affected stock markets.

riods, and this generally can be observed for all DAX options across all maturities and moneyness levels.[5]

Thus, in addition to the extension of the current research on certain DAX implied volatility patterns, this study also exhibits how DAX implied volatilities move during different market periods. This knowledge can be used by option traders and risk managers. Options traders should consider this risk factor by demanding a higher risk premium for option series that are highly exposed to remarkable changes of DAX implied volatilities. Additionally, the high time series dependencies among the implied volatilities across different options series in volatile periods reflect close price movements that should be analysed by risk managers in stress scenarios. Because put options are typically used by institutional investors for portfolio insurance, this risk factor exists for many investors.

Furthermore, the results provide evidence of regularities in the dynamic structure of the DAX IVS, which may be considered when predicting implied volatility. Gonçalves and Guidolin (2006), Konstantinidi et al. (2008), and Bernales and Guidolin (2014) also find evidence for predictable features of the IVS and propose corresponding approaches to exploit these regularities to generate better volatility forecasts. However, because the empirical findings in this study suggest that the considered DAX implied volatility series is non-stationary, the author does not pursue this modelling approach for DAX implied volatility, and leaves it for future research.[6]

Because the empirical analysis of the DAX IVS that is presented in Chapter 3 demonstrates that the constant volatility assumption of the BS model is violated, an option pricing model that is sufficiently flexible to allow for these features is necessary. Thus, in Chapter 4, four different classes of option pricing models are discussed, namely, stochastic volatility models, jump-diffusion models, local volatil-

[5] Interestingly, although the DAX level increased in 2009 and DAX implied volatilities for short-term options decreased, they did not return to their pre-crash levels.

[6] The interaction between non-stationarity, structural breaks and long memory effects generates a complex process structure. Perhaps future work on stochastic processes will provide solutions to model such structures.

ity models, and the concept of model-free implied volatility. Based on the models' ability to match the observed DAX IVS and the results of the empirical studies on the forecasting performance of implied volatility and the time series models that are provided in Chapter 5, the concept of model-free implied volatility that was developed by Britten-Jones and Neuberger (2000) is selected and used in this study to forecast DAX volatility. Furthermore, despite its well-documented weaknesses, the BS model is employed as a benchmark model in this study, because it can be regarded as a heuristic rule that is applied by many market participants.[7]

In addition to these option pricing models, the GARCH model, the EGARCH model, the ARFIMA model, and a modified form of the HAR model that was recently suggested by Corsi (2009) are employed because of their capabilities to reproduce the observed time series characteristics. Based on information criteria, the GARCH(1,2), the EGARCH(2,1), and the ARFIMA(1,d,1) models are selected to predict DAX volatility. Furthermore, the GARCH(1,1) model, which is used in many empirical applications as a benchmark model is employed to forecast DAX volatility. Besides the individual forecasts, this study also considers combined forecasts, because some studies provide evidence that forecast combinations can outperform individual forecasts.[8]

The generation and evaluation of the DAX volatility forecasts are described in Chapter 6. Because the data set contains long time series that cover clearly different volatility periods, this study investigates whether the high persistence in variance and long memory effects can be explained by structural breaks. Based on the Bai and Perron (2003) test, two breakpoints that can be linked to certain historical events are identified for the realised DAX volatility series.[9] The analysis suggests that the level of the long memory parameter for the considered realised DAX volatility series is partly driven by structural breaks. However, the long memory effect does not

[7] For a detailed explanation of the selection of the prediction models, see Chapter 5.4.

[8] See, e.g., the comprehensive studies by Makridakis and Hibon (2000), Stock and Watson (1999), and Marcellino (2004).

[9] The first breakdate marks the end of the volatile period at the beginning of the sample, which was driven by investors' fears of an impending recession in the US and the Iraq war in 2003. The second breakdate corresponds to the beginning of the financial crisis in 2008.

entirely disappear if structural breaks are removed from the series. As a result, the above ARFIMA model overestimates the long memory effect of the realised DAX volatility for the full sample. Nevertheless, the model captures an actual feature of the time series.

This issue has been neglected by the existing studies on the prediction of volatility that are based on long memory models (e.g., Lazarov (2004)). Particularly, all studies that compare the forecasting performance of implied volatility and time series models that are based on a longer time series likely suffer from this problem. Thus, the results that these studies provide must be interpreted with care, especially when they report superior prediction results in favour of implied volatility.

Because in this study, the prediction of DAX volatility is based on rolling windows of fixed sample sizes of 500 observations, a series of subsamples that are not affected by the observed structural breaks exists. Additionally, the empirical analysis demonstrates that long-range dependencies exist, even if structural breaks are removed from the realised volatility series. Therefore, the ARFIMA model is used to produce DAX volatility forecasts. Furthermore, the effects of unknown structural breaks on the modified HAR and GARCH model parameters are also examined. The test results suggest that no structural change emerged in the parameters of the modified HAR model and the GARCH models across the sample.

The DAX volatility forecasts are calculated based on the above mentioned models for different forecast horizons. The prediction results are evaluated by the MCS approach, where the MSE and the QLIKE are used as loss functions.[10] Because the sample period contains the 2008 financial crisis, the impact of the crisis on model forecasting performance is analysed by using the full sample and the subperiod that excludes the two most volatile months of the crisis.

The evaluation results regarding one-day-ahead DAX volatility forecasts confirm the findings of Lazarov (2004), who reports that (extended) ARFIMA models provide

[10] Patton (2011) demonstrates that both functions are robust to possible noise in the volatility proxy.

the best one-day-ahead DAX volatility forecasts. However, this study also extends this line of research. Similar to the findings of this study, Lazarov (2004) also reports that (extended) ARFIMA models provide better one-day-ahead DAX volatility forecasts than (extended) GARCH models. However, the evaluation approach that is employed by Lazarov (2004) does not allow for an analysis regarding whether the observed forecast error differences are significant. By applying the MCS approach, this work demonstrates that the ARFIMA and the related HAR model significantly outperform the GARCH and the EGARCH models. These results are robust to the employed loss function as well as the considered sample. Additionally, the above results show that the simple HAR model provides remarkable short-term forecasting performance for DAX volatility compared with a wide class of more sophisticated volatility prediction models. In total, these results demonstrate that the ability of the HAR and the ARFIMA models to capture long memory dependencies in realised DAX volatilities seem to provide useful information for predicting short-term DAX volatility.

Furthermore, this study refines the results of Claessen and Mittnik (2002) and Lazarov (2004) for two-weeks-ahead DAX volatility forecasts. Similar to Claessen and Mittnik (2002), this work shows that the combination of two individual forecasts that contain relevant information concerning future DAX volatility provides lower MSE prediction errors than individual forecasts. However, in contrast to Claessen and Mittnik (2002), who conclude that the combination of GARCH forecasts and BS ATM implied volatility seem to perform best, this study generally reports a significantly superior forecasting performance for the combined forecasts of BS ATM implied volatility and long memory models.

Lazarov (2004) evaluates the volatility forecasts based on the heteroscedasticity-consistent mean square error and also finds better prediction results for an ARFIMA model that is enhanced with implied volatility than for GARCH or extended GARCH models. However, Lazarov (2004) reports similar results for a linear regression model where the realised variance is regressed over the corresponding forecasting horizon on the lagged value of implied volatility.[11] Because of the employed evaluation method, Lazarov's (2004) results do not provide evidence on which approach performs significantly better than the other considered models. Based on the MCS approach and a longer sample period, this study shows that in general, the combined forecasts of BS ATM implied volatility and long memory models provide a significantly better forecasting performance than alternative models. Thus, similar to the results for one-day-ahead forecasts, DAX volatility predictions that are based on long memory models can improve BS ATM implied volatility forecasts by incorporating additional information.

The MCS results for one-month-ahead DAX volatility forecasts show that GARCH models provide significantly higher prediction errors than most other models that are under consideration. Thus, the results of this thesis are consistent with the findings of Raunig (2006) and Muzzioli (2010), who report a better forecasting performance for BS ATM implied volatility than for the GARCH models.[12]

In addition, the evaluation results show that BS ATM implied volatility provides a better forecasting performance than model-free volatility in both samples. These findings do not agree with Schöne (2010) and Tallau (2011), who report that the information content of the VDAX-New (or model-free implied volatility) is higher than the information content of the VDAX (or BS ATM implied volatility). Because the VDAX is based on a constant time to maturity of 45 days and the VDAX-New is calculated for a fixed maturity of 30 days, the better one-month-ahead forecasting performance of the VDAX-New is not surprising. In this study, model-free volatility is explicitly calculated for a maturity of one month, such that there is no mismatch between the maturities of the two implied volatility measures.

In fact, the results of this study confirm the findings of Muzzioli (2010), who also reports that BS implied volatility exhibits a better performance than model-free

[11] See Lazarov (2004), pp. 62.

[12] In contrast, Claessen and Mittnik (2002) find better prediction results of the GARCH(1,1) model compared with BS ATM implied volatility for one-month-ahead DAX volatility predictions, but they consider only a small sample of 26 four-week forecasts. See Claessen and Mittnik (2002), p. 313.

volatility. However, Muzzioli (2010) does not consider DAX volatility forecasts based on long memory models and combined forecasts. Overall, the above findings suggest that these forecasting approaches provide superior prediction results.

In summary, the empirical results that are presented by this thesis contribute to the current research in several ways. First, the existing studies mostly perform partial model comparisons that ignore certain aspects. In particular, volatility forecasts that are based on long memory models and model-free implied volatility are not considered in the same forecast comparison regarding DAX volatility. This study demonstrates that forecasts based on long memory models typically provide important information to predict DAX volatility, either as individual predictions or combined with other prediction models. Interestingly, this result does not depend on the forecast horizon or the loss function that is considered. If DAX volatility is predicted for longer horizons (from two weeks to one month), the information that DAX implied volatility provides should be considered because it supplements the information that is contained in DAX realised volatility.

Second, past studies of the German stock market do not take recently developed forecast evaluation approaches into account. Most studies use encompassing regressions that consider individual forecast comparisons and do not control for data snooping effects to evaluate volatility forecasts. Third, squared daily returns that provide a noisy estimate of latent volatility are often used as a volatility proxy to evaluate DAX volatility predictions. However, Hansen and Lunde (2006) demonstrated that a ranking of volatility forecasting models that is based on the realised variance is more likely to be consistent with the true ranking.[13] Fourth, although some studies detect long memory effects in financial time series, the effects of structural breaks on the applied forecasting methodology are not examined. By considering all of these issues, this thesis provides further information to investment and risk managers regarding which forecasting method delivers superior DAX volatility forecasts.

[13]See Hansen and Lunde (2006), pp. 98-100.

Additional studies are recommended to extend these findings. As mentioned, the relatively poor performance of one-day-ahead volatility forecasts that are based on implied volatility is not surprising because it can be explained by the mismatch between the forecast horizon and the maturity. It would be interesting to determine whether short-term options (e.g., DAX weekly options) contain relevant information for short-term volatility forecasts. Further, the extension of the ARFIMA model by a Monday dummy variable shows, that taking calendar day effects into account can enhance model performance. Therefore, future research should examine, whether such variables provide incremental information for predicting volatility (e.g., calendar day effects, seasonalities, and macroeconomic announcements).

In addition, the systematic dynamic pattern of the DAX IVS can be used by time series models to enhance the forecasting performance of implied volatility. Moreover, recent studies by Hansen et al. (2011a) as well as Louzis et al. (2014) provide evidence, that the Realized GARCH model structure can lead to a better empirical fit over standard GARCH models, respectively, generate superior Value-at-Risk estimates. Maybe this new framework can help to improve GARCH volatility forecasts, and provide new insights into modeling and forecasting intraday volatility.

Another point left for future research is the construction of combined volatility forecasts. Because structural breaks are difficult to detect in the data generating process, combined forecasts that are based on models that adapt to these changes differently can provide better forecasts on average and have particular interest in unstable markets.[14] While this study combines individual volatility forecast-based equal weights, alternative methods to derive “optimal” weights can be employed. Additionally, the above analysis concerning the effects of structural breaks on the applied forecasting methodology can only provide initial insights that should be further investigated. Finally, future research should also examine whether and how alternative forecast evaluation methods influence the forecast rankings that have been presented.

[14] For an explanation, see Pesaran and Timmermann (2007).

A. Appendix of Chapter 3

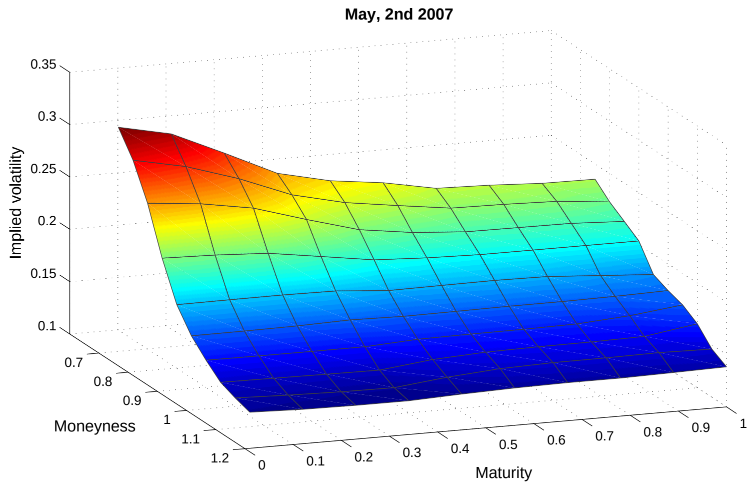


Figure A.1.: DAX implied volatility surface on May, 2nd 2007.

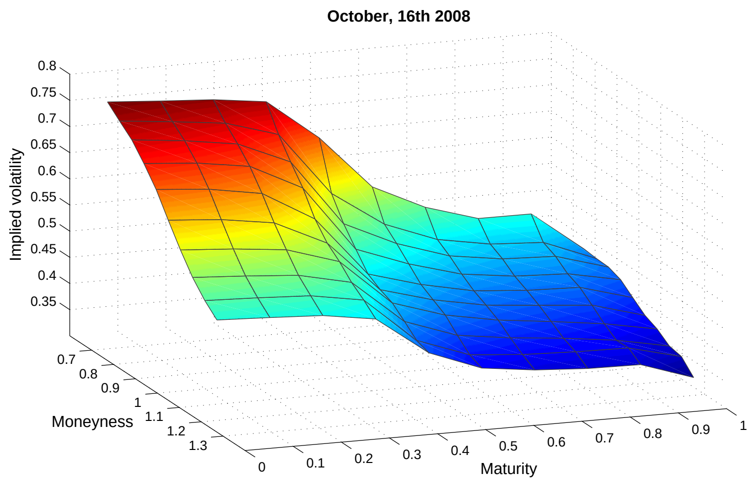


Figure A.2.: DAX implied volatility surface on October, 16th 2008.

B. Appendix of Chapter 6

Table B.1.: ADF test results for the null hypothesis “random walk with drift”

	ADF (10)	ADF (20)
mfv	-2.46	-2.37
	0.7%	0.9%
lnmfv	-2.30	-2.15
	1.1%	1.6%
bsatm	-2.63	-2.63
	0.4%	0.4%
lnbsatm	-2.38	-2.28
	0.9%	1.1%
rvola	-3.76	-3.23
	0.0%	0.1%
lnrvola	-3.48	-2.83
	0.0%	0.2%
rdax	-13.63	-10.58
	0.0%	0.0%

Source: own calculations.

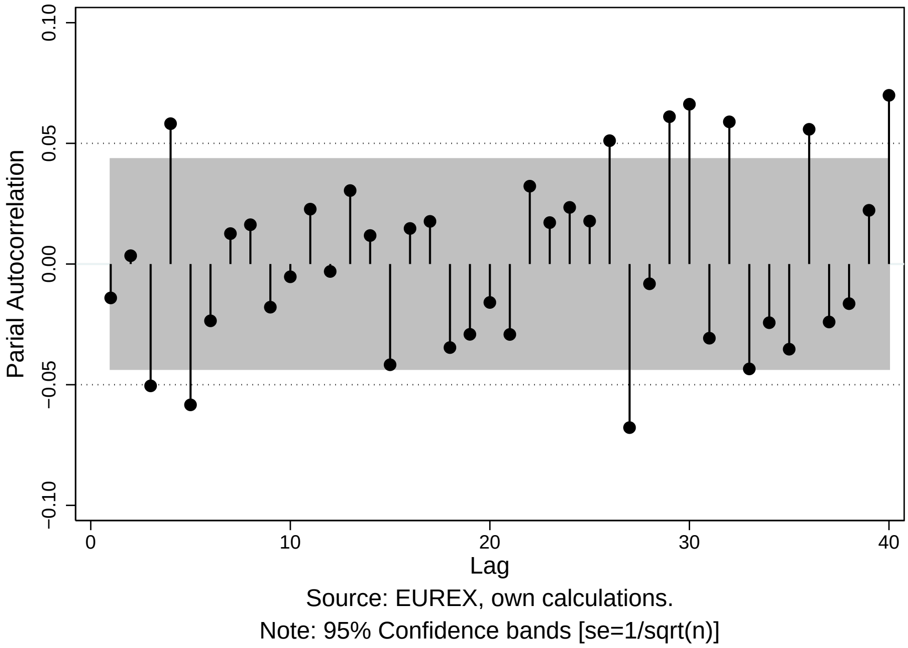


Figure B.1.: Partial autocorrelation function for DAX return series

Table B.2.: Information criteria for GARCH model selection

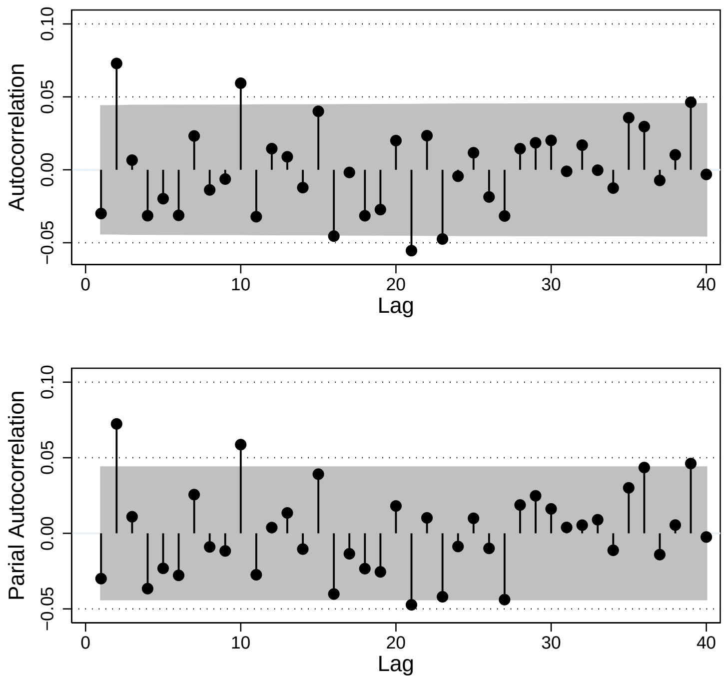
model	AIC	SIC	HSIC
ARCH(1)	-11,490.7	-11,468.2	-16,937.9
ARCH(2)	-11,651.1	-11,623.0	-17,221.7
GARCH(1,1)	-11,963.8	-11,935.8	-17,678.7
GARCH(1,2)	-11,970.2	-11,936.5	-17,682.1
GARCH(2,1)	-11,925.7	-11,892.0	-17,605.7
GARCH(2,2)	-11,926.8	-11,887.5	-17,603.5
EARCH(1)	-11,481.0	-11,452.9	-16,907.5
EARCH(2)	-11,582.0	-11,542.7	-17,083.5
EGARCH(1,1)	-12,022.6	-11,988.9	-17,747.6
EGARCH(1,2)	-12,021.5	-11,982.2	-17,741.5
EGARCH(2,1)	-12,045.3	-12,000.3	-17,759.4
EGARCH(2,2)	-12,048.5	-11,997.9	-17,758.1

Source: own calculations.

Table B.3.: Information criteria for ARFIMA model selection

model	AIC	SIC
ARFIMA(1,d,0)	84.9	107.4
ARFIMA(0,d,1)	85.3	107.8
ARFIMA(1,d,1)	64.5	92.6
ARFIMA(2,d,0)	85.7	113.8
ARFIMA(0,d,2)	85.7	113.7
ARFIMA(2,d,1)	65.7	99.4
ARFIMA(1,d,2)	65.8	99.5
ARFIMA(2,d,2)	67.3	106.7

Source: own calculations.



Source: EUREX, own calculations.

Note: Bartlett's formula for MA(q) 95% confidence bands (top panel) and 95% confidence bands [se=1/sqrt(n)] (bottoom panel).

Figure B.2.: Correlograms of HAR model residuals

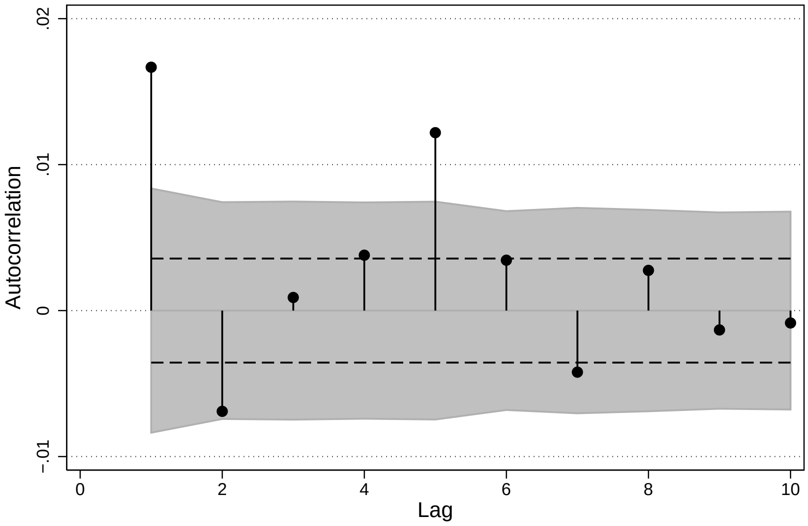


Figure B.3.: Correlogram of DAX 5-minute returns

Table B.4.: Estimation results for an MA(2) model fitted to DAX returns

	MA(2)
μ	4.95 e-07
	(2.52 e-06)
ψ_1	0.017***
	(0.004)
ψ_5	0.012***
	(0.004)
Q1*(10)	7.61
	66.68%

Source: own calculations.
Note: Standard error in parentheses;
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.
The results of Diebold's ARCH robust Q-statistic are denoted by Q1*.

Table B.5.: ADF test results for one-day loss differentials

	garch11	egarch21	arfima11	har	bsatm	mfv
egarch21	-6.25					
	0.00%					
arfima11	-5.70	-5.74				
	0.00%	0.00%				
har	-6.45	-5.54	-6.93			
	0.00%	0.00%	0.00%			
bsatm	-5.37	-5.68	-8.36	-7.34		
	0.00%	0.00%	0.00%	0.00%		
mfv	-5.33	-5.88	-7.76	-7.23	-6.78	
	0.00%	0.00%	0.00%	0.00%	0.00%	
rw	-11.08	-8.29	-9.09	-9.86	-7.87	-7.51
	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%

Source: own calculations.

Table B.6.: ADF test results for two-weeks loss differentials

	garch11	egarch21	arfima11	har	bsatm	mfv
egarch21	-6.75					
	0.00%					
arfima11	-8.02	-4.70				
	0.00%	0.01%				
har	-6.25	-4.80	-8.22			
	0.00%	0.01%	0.00%			
bsatm	-5.29	-4.59	-4.64	-4.22		
	0.00%	0.01%	0.01%	0.06%		
mfv	-6.79	-4.43	-4.26	-4.26	-6.04	
	0.00%	0.03%	0.05%	0.05%	0.00%	
rw	-7.99	-5.79	-6.57	-6.15	-5.21	-6.14
	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%

Source: own calculations.

Table B.7.: ADF test results for one-month loss differentials

	garch11	egarch21	arfima11	har	bsatm	mfv
egarch21	-4.34					
	0.04%					
arfima11	-6.12	-3.87				
	0.00%	0.22%				
har	-6.12	-4.09	-5.38			
	0.00%	0.10%	0.00%			
bsatm	-3.61	-3.95	-4.27	-3.97		
	0.56%	0.17%	0.05%	0.16%		
mfv	-4.37	-3.83	-4.08	-3.63	-5.19	
	0.03%	0.26%	0.10%	0.53%	0.00%	
rw	-5.12	-4.50	-5.96	-6.16	-7.26	-7.50
	0.00%	0.02%	0.00%	0.00%	0.00%	0.00%

Source: own calculations.

Table B.8.: Estimation results for AR(p) processes of one-day loss differentials (1/2)

loss series	δ	α_1	α_2	α_3	α_4	α_5	α_6	α_7
d12	0.0004	0.4589***	-0.1604***	-0.1022***	0.0183	0.1755***	-0.0819***	0.0794***
	(0.0005)	(0.0069)	(0.0067)	(0.0105)	(0.0125)	(0.0154)	(0.0157)	(0.0159)
d13	0.0012*	0.2675***	0.0708***	-0.1293***	-0.0047	0.1020***	0.0607***	0.0188*
	(0.0007)	(0.0075)	(0.0089)	(0.0089)	(0.0165)	(0.0178)	(0.0172)	(0.0098)
d14	0.0011	0.2495***	0.0157	-0.1436***	-0.0629***	0.0968***	0.0599***	-0.0154
	(0.0007)	(0.0097)	(0.0100)	(0.0078)	(0.0105)	(0.0139)	(0.0161)	(0.0096)
d15	0.0005	0.2599***	0.0661***	-0.0459***	-0.0089	0.0819***	0.1175***	0.0106
	(0.0008)	(0.0075)	(0.0078)	(0.0082)	(0.0096)	(0.0125)	(0.0163)	(0.0170)
d16	-0.0002	0.2721***	0.0568***	-0.0586***	-0.0264***	0.0940***	0.1246***	-0.0062
	(0.0008)	(0.0069)	(0.0082)	(0.0087)	(0.0093)	(0.0129)	(0.0168)	(0.0178)
d17	0.0000	0.1216***	-0.0340***	-0.1016***	-0.0879***	0.0333*	0.0059	0.1288***
	(0.0005)	(0.0114)	(0.0128)	(0.0131)	(0.0115)	(0.0189)	(0.0133)	(0.0051)
d23	0.0007***	0.2070***						
	(0.0002)	(0.0126)						
d24	0.0007	0.1011***	0.0375***	0.0495***	0.0630***	0.0286**	-0.0167	-0.0368
	(0.0005)	(0.0120)	(0.0101)	(0.0140)	(0.0191)	(0.0116)	(0.0135)	(0.0233)
d25	0.0000	0.2274***	-0.0159*	-0.0169	0.0758***	0.1248***	0.0447***	-0.0271
	(0.0004)	(0.0145)	(0.0092)	(0.0136)	(0.0141)	(0.0132)	(0.0158)	(0.0175)
d26	-0.0006	0.1739***	0.0274**	0.0228	0.0651***	0.1175***	0.0547***	-0.0516**
	(0.0004)	(0.0180)	(0.0131)	(0.0142)	(0.0151)	(0.0150)	(0.0195)	(0.0212)
d27	-0.0005	0.0185	-0.0268*	-0.0566***	-0.0572***	-0.0394*	0.0036	0.2601***
	(0.0006)	(0.0121)	(0.0143)	(0.0165)	(0.0077)	(0.0227)	(0.0147)	(0.0081)
d34	-0.0001	0.0947***	0.1781***	-0.0524***	-0.0195	0.0466***	0.0549***	0.1362***
	(0.0002)	(0.0122)	(0.0100)	(0.0135)	(0.0180)	(0.0100)	(0.0129)	(0.0102)
d35	-0.0007***	0.1170***						
	(0.0001)	(0.0122)						
d36	-0.0013***	0.1410***						
	(0.0002)	(0.0137)						
d37	-0.0008***	-0.0774***	-0.1067***					
	(0.0002)	(0.0101)	(0.0095)					
d45	-0.0006***	0.1290***	0.0983***	0.0747***				
	(0.0002)	(0.0121)	(0.0094)	(0.0131)				
d46	-0.0013***	0.1323***	0.1134***	0.0831***				
	(0.0003)	(0.0154)	(0.0096)	(0.0114)				
d47	-0.0007***	-0.0323***	-0.0761***					
	(0.0002)	(0.0108)	(0.0134)					
d56	-0.0007***	0.2695***	0.0716***	-0.0212	-0.0362**	0.0770***	0.0302*	0.0290
	(0.0001)	(0.0140)	(0.0122)	(0.0152)	(0.0161)	(0.0133)	(0.0175)	(0.0182)
d57	-0.0005	0.1197***	-0.0005	0.0572**	-0.0393*	0.0505***	0.0240**	0.2618***
	(0.0008)	(0.0113)	(0.0187)	(0.0283)	(0.0235)	(0.0187)	(0.0102)	(0.0049)
d67	0.0001	0.0930***	0.0025	0.0727***	-0.0488**	0.0517***	0.0150	0.2813***
	(0.0008)	(0.0123)	(0.0174)	(0.0231)***	(0.0215)	(0.0166)	(0.0148)	(0.0048)

Source: own calculations.

Note: Standard error in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. The parameters δ, α_1, α_2,..., α_p refer to an AR(p) process of the form $x_t = \delta + \alpha_1 x_{t-1} + \alpha_2 x_{t-2} + \ldots + \alpha_p x_{t-p} + \varepsilon_t$. The expression d_{ij} represents the loss differentials which are calculated from the individual loss series of the forecasting models. The index denotes the employed forecasting models which are coded by 1 = GARCH(1,1), 2 = EGARCH(2,1), 3 = ARFIMA(1,d,1), 4 = HAR, 5 = BSATM, 6 = MFV, and 7 = RW.

Table B.9.: Estimation results for AR(p) processes of one-day loss differentials (2/2)

loss series	α_8	α_9	α_{10}	α_{11}	α_{12}	α_{13}	α_{14}	α_{15}	BIC
d12	0.0985***								-10,546
	(0.0134)								
d13	0.1418***								-10,723
	(0.0117)								
d14	0.1222***	0.0748***	-0.0066	-0.0038	0.1512***	-0.1095***	0.1550***		-10,799
	(0.0116)	(0.0135)	(0.0102)	(0.0098)	(0.0111)	(0.0097)	(0.0097)		
d15	0.0637***	0.1084***							-10,665
	(0.0158)	(0.0146)							
d16	0.0454***	0.0988***							-10,327
	(0.0145)	(0.0139)							
d17	-0.1146***								-8,513
	(0.0105)								
d23									-11,544
d24	0.0561***	0.0208*	0.0549***	0.1365***	0.0898***				-11,481
	(0.0102)	(0.0121)	(0.0116)	(0.0143)	(0.0156)				
d25	0.1692***								-11,295
	(0.0094)								
d26	0.1787***								-11,243
	(0.0096)								
d27									-8,786
d34	-0.1126***	-0.0398***	-0.1206***	-0.0074	0.2352***	-0.0607***	-0.0265***	0.1999***	-14,114
	(0.0079)	(0.0118)	(0.0113)	(0.0289)	(0.0063)	(0.0130)	(0.0088)	(0.0105)	
d35									-12,043
d36									-11,580
d37									-10,804
d45									-11,853
d46									-11,319
d47									-10,914
d56	-0.1681***	-0.0222	0.0198	0.0008	0.0965***	0.0153	0.0224	0.1621***	-15,138
	(0.0102)	(0.0172)	(0.0165)	(0.0173)	(0.0176)	(0.0142)	(0.0172)	(0.0143)	
d57	-0.1039***								-8,867
	(0.0261)								
d67	-0.0939***								-8,821
	(0.0269)								

Source: own calculations.
Note: Standard error in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. The parameters δ, α_1, α_2,..., α_p refer to an AR(p) process of the form $x_t = \delta + \alpha_1 x_{t-1} + \alpha_2 x_{t-2} + \ldots + \alpha_p x_{t-p} + \varepsilon_t$. The expression d_{ij} represents the loss differentials which are calculated from the individual loss series of the forecasting models. The index denotes the employed forecasting models which are coded by 1 = GARCH(1,1), 2 = EGARCH(2,1), 3 = ARFIMA(1,d,1), 4 = HAR, 5 = BSATM, 6 = MFV, and 7 = RW.

Table B.10.: Estimation results for AR(p) processes of two-weeks loss differentials

loss series	δ	α_1	α_2	α_3	α_4	BIC
d12	-0.0013	0.0992**				-772.0
	(0.0032)	(0.0414)				
d13	-0.0001					-857.2
	(0.0012)					
d14	-0.0001	0.3218***	-0.1921***	-0.1548**	-0.1953***	-927.1
	(0.0019)	(0.1010)	(0.0638)	(0.0710)	(0.0525)	
d15	0.0010	-0.1999***				-908.1
	(0.0024)	(0.0559)				
d16	0.0002	-0.3190***				-865.4
	(0.0022)	(0.0664)				
d17	0.0005	0.0888*	-0.4259***	-0.3072***		-982.2
	(0.0006)	(0.0496)	(0.0211)	(0.0300)		
d23	0.0012	0.5814***				-1,199.9
	(0.0015)	(0.0201)				
d24	0.0012	-0.1735***				-968.2
	(0.0014)	(0.0256)				
d25	0.0022	0.4894***				-890.0
	(0.0049)	(0.0284)				
d26	0.0014	0.5889***				-976.2
	(0.0046)	(0.0628)				
d27	0.0018	0.0682***				-735.3
	(0.0043)	(0.0251)				
d34	-0.0001	-0.6244***	-0.3144***			-1,139.5
	(0.0003)	(0.0216)	(0.0259)			
d35	0.0010	0.2903***				-974.9
	(0.0025)	(0.0321)				
d36	0.0002	0.2024	0.1957*			-1,043.4
	(0.0022)	(0.1439)	(0.1186)			
d37	0.0005	-0.0773***				-804.8
	(0.0032)	(0.0295)				
d45	0.0010	0.2546***	0.3582***			-1,035.8
	(0.0035)	(0.0796)	(0.0380)			
d46	0.0002	-0.0307	0.4528***			-1,008.8
	(0.0024)	(0.0450)	(0.0226)			
d47	0.0006	0.2732***				-924.9
	(0.0029)	(0.0232)				
d56	-0.0008	-0.2323***				-1,244.5
	(0.0008)	(0.0343)				
d57	-0.0005	-0.2501***				-939.6
	(0.0008)	(0.0177)				
d67	0.0002	-0.3054***				-856.6
	(0.0012)	(0.0192)				

Source: own calculations.
Note: Standard error in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.
The parameters δ, α_1, α_2,. . ., α_p refer to an AR(p) process of the form $x_t = \delta + \alpha_1 x_{t-1} + \alpha_2 x_{t-2} + \ldots + \alpha_p x_{t-p} + \varepsilon_t$. The expression d_{ij} represents the loss differentials which are calculated from the individual loss series of the forecasting models. The index denotes the employed forecasting models which are coded by 1 = GARCH(1,1), 2 = EGARCH(2,1), 3 = ARFIMA(1,d,1),4 = HAR, 5 = BSATM, 6 = MFV, and 7 = RW.

Table B.11.: Estimation results for AR(p) processes of one-month loss differentials

loss series	δ	α_1	α_2	BIC
d12	-0.0033			-318.5
	(0.0031)			
d13	0.0001	-0.3236***		-423.2
	(0.0021)	(0.0527)		
d14	0.0007	-0.3019***		-479.7
	(0.0011)	(0.0568)		
d15	0.0013	-0.3944***		-486.8
	(0.0018)	(0.0562)		
d16	0.0009	-0.5113***		-476.2
	(0.0011)	(0.0313)		
d17	0.0011			-512.7
	(0.0008)			
d23	0.0034	0.4104***		-393.0
	(0.0086)	(0.0606)		
d24	0.0040	0.2811***		-349.2
	(0.0119)	(0.0761)		
d25	0.0045	0.4420***		-334.8
	(0.0123)	(0.0503)		
d26	0.0040	0.4908***		-349.1
	(0.0119)	(0.0481)		
d27	0.0044	0.1373**		-290.8
	(0.0113)	(0.0593)		
d34	0.0006			-549.8
	(0.0006)			
d35	0.0012	0.3822***		-463.6
	(0.0033)	(0.0491)		
d36	0.0007	0.4203***		-487.5
	(0.0026)	(0.0573)		
d37	0.0009			-382.2
	(0.0020)			
d45	0.0006			-512.0
	(0.0008)			
d46	0.0002			-520.3
	(0.0008)			
d47	0.0004			-420.1
	(0.0015)			
d56	-0.0004			-672.2
	(0.0003)			
d57	-0.0003	-0.3898***	-0.2894***	-457.4
	(0.0008)	(0.0457)	(0.0581)	
d67	0.0001	-0.4345***	-0.3047***	-436.8
	(0.0009)	(0.0429)	(0.0511)	

Source: own calculations.
Note: Standard error in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. The parameters δ, α_1, and α_2 refer to an AR(p) process of the form $x_t = \delta + \alpha_1 x_{t-1} + \alpha_2 x_{t-2} + \ldots + \alpha_p x_{t-p} + \varepsilon_t$. The expression d_{ij} represents the loss differentials which are calculated from the individual loss series of the forecasting models. The index denotes the employed forecasting models which are coded by 1 = GARCH(1,1), 2 = EGARCH(2,1), 3 = ARFIMA(1,d,1), 4 = HAR, 5 = BSATM, 6 = MFV, and 7 = RW.

Bibliography

Acworth, W. (2014), FIA Annual Volume Survey: Commodity and Interest Rate Trading Push Trading Higher in 2013, *Futures Industry Magazine* pp. 14–26.

Äijö, J. (2008), Implied Volatility Term Structure Linkages between VDAX, VSMI and VSTOXX Volatility Indices, *Global Finance Journal* 18(3), pp. 290–302.

Aït-Sahalia, Y. (2002), Telling from Discrete Data Whether the Underlying Continuous-Time Model Is a Diffusion, *The Journal of Finance* 57(5), pp. 2075–2112.

Aït-Sahalia, Y., Bickel, P. J., and Stoker, T. M. (2001), Goodness-of-Fit Tests for Kernel Regression with an Application to Option Implied Volatilities, *Journal of Econometrics* 105(2), pp. 363–412.

Aït-Sahalia, Y. and Duarte, J. (2003), Nonparametric Option Pricing Under Shape Restrictions, *Journal of Econometrics* 116(1), pp. 9–47.

Aït-Sahalia, Y. and Lo, A. W. (1998), Nonparametric Estimation of State-Price Densities Implicit in Financial Asset Prices, *The Journal of Finance* 53(2), pp. 499–547.

Aït-Sahalia, Y. and Lo, A. W. (2000), Nonparametric Risk management and Implied Risk Aversion, *Journal of Econometrics* 94(1), pp. 9–51.

Aït-Sahalia, Y. and Mancini, L. (2008), Out of Sample Forecasts of Quadratic Variation, *Journal of Econometrics* 147(1), pp. 17–33.

Aït-Sahalia, Y., Mykland, P. A., and Zhang, L. (2011), Ultra High Frequency Volatility Estimation with Dependent Microstructure Noise, *Journal of Econometrics* 160(1), pp. 160–175.

Aït-Sahalia, Y., Wang, Y., and Yared, F. (2001), Do Option Markets Correctly Price the Probabilities of Movement of the Underlying Asset?, *Journal of Econometrics* 102(1), pp. 67–110.

Alexander, C. (2008), *Pricing, Hedging and Trading Financial Instruments*, Chichester, UK: John Wiley & Sons.

Alexander, C. and Nogueira, L. M. (2004a), Hedging with Stochastic and Local Volatility. ISMA Centre Discussion Paper No. DP2004-11, The University of Reading.

Alexander, C. and Nogueira, L. M. (2004b), Stochastic Local Volatility. Working Paper, ICMA Centre, The Universtiy of Reading.

Andersen, A. B. and Wagener, T. (2002), Extracting Risk Neutral Probability Densities by Fitting Implied Volatility Smiles: Some Methodological Points and an Application to the 3M Euribor Futures Options Prices. ECB Working Paper No. 198, European Central Bank.

Andersen, L. B. G. and Brotherton-Ratcliffe, R. (2000), The Equity Option Volatility Smile: An Implicit Finite-Difference Approach, *Journal of Computational Finance* 3(2), pp. 5–37.

Andersen, T. G. and Benzoni, L. (2009), Realized Volatility, in T. Mikosch, J.-P. Kreiß, R. A. David, and T. G. Andersen (eds.), *Handbook of Financial Time Series*, Berlin: Springer, pp. 555–575.

Andersen, T. G. and Bollerslev, T. (1998), Answering the Skeptics: Yes, Standard Volatility Models Do Provide Accurate Forecasts, *International Economic Review* 39(4), pp. 885–905.

Andersen, T. G., Bollerslev, T., and Diebold, F. X. (2007), Roughing It Up: Including Jump Components in the Measurement, Modeling and Forecasting of Return Volatility, *Review of Economics and Statistics* 89(4), p. 701–720.

Andersen, T. G., Bollerslev, T., and Meddahi, N. (2011), Realized Volatility Forecasting and Market Microstructure Noise, *Journal of Econometrics* 160(1), pp. 220–234.

Andersen, T. G. et al. (2000), Exchange Rate Returns Standardized by Realized Volatility Are (Nearly) Gaussian, *Multinational Finance Journal* 4(3&4), pp. 159–179.

Andersen, T. G. et al. (2001a), The Distribution of Exchange Rate Volatility, *Journal of the American Statistical Association* 96(453), pp. 42–55.

Andersen, T. G. et al. (2001b), The Distribution of Realized Stock Return Volatility, *Journal of Financial Economics* 61(1), pp. 43–76.

Andersen, T. G. et al. (2003), Modeling and Forecasting Realized Volatility, *Econometrica* 71(2), pp. 579–625.

Andersen, T. G. et al. (2006), Volatility and Correlation Forecasting, in G. Elliott, C. W. J. Granger, and A. Timmermann (eds.), *Handbook of Economic Forecasting*, Amsterdam: Elsevier, vol. 1, pp. 779–865.

Andrews, D. W. K. (1993), Tests for Parameter Instability and Structural Change with Unknown Change Point, *Econometrica* 61(4), pp. 821–856.

Andrews, D. W. K. and Ploberger, W. (1994), Optimal Tests When a Nuisance Parameter is Present Only Under the Alternative, *Econometrica* 62(6), pp. 1383–1414.

Angelidis, T., Benos, A., and Degiannakis, S. (2004), The Use of GARCH models in VaR Estimation, *Statistical Methodology* 1(1-2), pp. 105–128.

Ashley, R. A. and Patterson, D. M. (2010), Apparent Long Memory in Time Series As an Artifact of a Time-Varying Mean: Considering Alternatives to the Fractionally Integrated Model, *Macroeconomic Dynamics* 14(S1), pp. 59–87.

Audrino, F. and Hu, Y. (2011), Volatility Forecasting: Downside Risk, Jumps and Leverage Effect. Discussion Paper no. 2011-38, Universität St. Gallen.

Avellaneda, M. et al. (1997), Calibrating Volatility Surfaces via Relative-Entropy Minimization, *Applied Mathematical Finance* 4(1), pp. 37–64.

Awartani, B. M. A. and Corradi, V. (2005), Predicting the Volatility of the S&P-500 Stock Index via GARCH Models: The Role of Asymmetrics, *International Journal of Forecasting* 21(1), pp. 167–183.

Back, K. (1992), Insider Trading in Continuous Time, *Review of Financial Studies* 5(3), pp. 387–410.

Back, K. (1993), Asymmetric Information and Options, *The Review of Financial Studies* 6(3), pp. 435–472.

Backus, D. et al. (2004), Accounting for Biases in Black-Scholes. Working Paper, New York University.

Bai, J. and Perron, P. (1998), Estimating and Testing Linear Models with Multiple Structural Changes, *Econometrica* 66(1), pp. 47–78.

Bai, J. and Perron, P. (2003), Computation and Analysis of Multiple Structural Change Models, *Journal of Applied Econometrics* 18(1), pp. 1–22.

Baillie, R. and Bollerslev, T. (1989), Common Stochastic Trends in a System of Exchange Rates, *Journal of Monetary Economics* 44(1), pp. 167–181.

Baillie, R. T., Bollerslev, T., and Mikkelsen, H. O. (1996), Fractionally Integrated Generalized Autoregressive Conditional Heteroskedasticity, *Journal of Econometrics* 74(1), pp. 3–30.

Bakshi, G., Cao, C., and Chen, Z. (1997), Empirical Performance of Alternative Option Pricing Models, *The Journal of Finance* 52(5), pp. 2003–2049.

Bandi, F. M. and Russell, J. R. (2006), Separating Microstructure Noise from Volatility, *Journal of Financial Economics* 79(3), pp. 655–692.

Banerjee, A. and Urga, G. (2005), Modelling Structural Breaks, Long Memory and Stock Market Volatility: An Overview, *Journal of Econometrics* 129(1-2), pp. 1–34.

Barndorff-Nielsen, O. et al. (2008), Designing Realized Kernels to Measure the Ex-Post Variation of Equity Prices in the Presence of Noise, *Econometrica* 76(6), pp. 1481–1536.

Barndorff-Nielsen, O. E. and Shephard, N. (2004), Power and Bipower Variation with Stochastic Volatility and Jumps, *Journal of Financial Econometrics* 2(1), pp. 1–37.

Barndorff-Nielsen, O. E. and Shephard, N. (2007), Variation, Jumps, Market Frictions and High Frequency Data in Financial Econometrics, in R. Blundell, W. Newey, and T. Persson (eds.), *Advances in Economics and Econometrics*, Cambridge, UK: Cambridge UP, vol. 3, pp. 328–372.

Barone-Adesi, G. and Whaley, R. E. (1987), Efficient Analytic Approximation of American Option Values, *The Journal of Finance* 42(2), pp. 301–320.

Bates, D. S. (1991), The Crash of '87: Was It Expected? The Evidence from Options Markets, *The Journal of Finance* 46(3), pp. 1009–1044.

Bates, D. S. (1996a), Dollar Jump Fears, 1984-1992: Distributional Abnormalities Implicit in Currency Futures Options, *Journal of International Money and Finance* 15(1), pp. 65–93.

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

Bates, D. S. (2000), Post-'87 Crash Fears in the S&P 500 Futures Option Market, *Journal of Econometrics* 94(1-2), pp. 181–238.

Baum, C. F. (2000), Tests for Long Memory in a Time Series, *Stata Technical Bulletin* 10(57), pp. 39–44.

Bauwens, L., Hafner, C. M., and Laurent, S. (2012), Volatility Models, in L. Bauwens, C. M. Hafner, and S. Laurent (eds.), *Handbook of Volatility Models and Their Applications*, Hoboken, NJ: John Wiley & Sons, vol. 3, pp. 1–48.

Beaglehole, D. and Chebanier, A. (2002), Mean Reversion with a Smile, *Risk* 15(4), pp. 95–98.

Becker, R. and Clements, A. E. (2008), Are Combination Forecasts of S&P 500 Volatility Statistically Superior?, *International Journal of Forecasting* 24(1), pp. 122–133.

Becker, R., Clements, A. E., and White, S. I. (2006), On the Informational Efficiency of S&P 500 Implied Volatility, *North American Journal of Economics and Finance* 17(2), pp. 139–153.

Becker, R., Clements, A. E., and White, S. I. (2007), Does Implied Volatility Provide any Information Beyond that Captured in Model-Based Volatility Forecasts?, *Journal of Banking & Finance* 31(8), pp. 2535–2549.

Beckers, S. (1981), Standard Deviations Implied in Option Prices as Predictors of Future Stock Price Volatility, *Journal of Banking & Finance* 5(3), pp. 363–382.

Behr, A. (2005), *Einführung in die Statistik mit R*, München: Vahlen.

Beinert, M. and Trautmann, S. (1995), Verlaufsmuster der impliziten Aktienvolatilität – Beobachtungen am Frankfurter Optionsmarkt und der Deutschen Terminbörse. Working Paper, Johannes-Gutenberg-Universität Mainz.

Bekaert, G., Hoerova, M., and Duca, M. L. (2013), Risk, Uncertainty and Monetary Policy, *Journal of Monetary Economics* 60(7), pp. 771–788.

Benko, M., Härdle, W., and Kneip, A. (2009), Common Functional Principal Components, *The Annals of Statistics* 37(1), pp. 1–34.

Benzoni, L., Collin-Dufresne, P., and Goldstein, R. S. (2011), Explaining Asset Pricing Puzzles Associated with the 1987 Market Crash, *Journal of Financial Economics* 101(3), pp. 552–573.

Beran, J. (1994), *Statistics for Long-Memory Processes*, Boca Raton, FL: Chapman & Hall.

Berkowitz, J. (2010), On Justifications for the ad hoc Black-Scholes Method of Option Pricing, *Studies in Nonlinear Dynamics & Econometrics* 14(1), pp. 96–120.

Bernales, A. and Guidolin, M. (2014), Can We Forecast the Implied Volatility Surface Dynamics of Equity Options? Predictability and Economic Value Tests, *Journal of Banking & Finance* 46, pp. 326–342.

Bertsch, M. (2008), *Das Phänomen der Volatilität: Die Beurteilung der Informationseffizienz des europäischen Optionsmarktes*, Berlin: dissertation.de.

Birke, M. and Pilz, K. F. (2009), Nonparametric Option Pricing with No-Arbitrage Constraints, *Journal of Financial Econometrics* 7(2), pp. 53–76.

Birru, J. and Figlewski, S. (2012), Anatomy of a Meltdown: The Risk Neutral Density for the S&P 500 in the Fall of 2008, *Journal of Financial Markets* 15(2), pp. 151–180.

Black, F. (1975), Fact and Fantasy in the Use of Options, *Financial Analysts Journal* 31(4), pp. 36–72.

Black, F. (1976), Studies of Stock Market Volatility Changes, *Proceedings of the American Statistical Association, Business and Economic Statistics Section* p. 177–181.

Black, F. and Scholes, M. (1973), The Pricing of Options and Corporate Liabilities, *Journal of Political Economy* 81(3), pp. 637–659.

Blair, B. J., Poon, S.-H., and Taylor, S. J. (2001), Forecasting S&P 100 Volatility: The Incremental Information Content of Implied Volatilities and High-Frequency Index Returns, *Journal of Econometrics* 105(1), pp. 5–26.

Bliss, R. R. and Panigirtzoglou, N. (2002), Testing the Stability of Implied Probability Density Functions, *Journal of Banking & Finance* 26(2-3), pp. 381–422.

Bloom, N. (2014), Fluctuations in Uncertainty, *Journal of Economic Perspectives* 28(2), pp. 153–176.

Bluhm, H. H. W. and Yu, J. (2001), Forecasting Volatility: Evidence from the German Stock Market. Department of Economics Working Paper Series 219, The University of Auckland.

Bolek, A. (1999), *Volatilitätsschwankungen und DAX-Optionen: Auswirkungen auf Bewertung und Risikomanagement*, Wiesbaden: Deutscher Universitäts-Verlag.

Bollen, N. P. B. and Whaley, R. E. (2004), Does Net Buying Pressure Affect the Shape of Implied Volatility Functions?, *The Journal of Finance* 59(2), pp. 711–753.

Bollerslev, T. (1986), Generalized Autoregressive Conditional Heteroskedasticity, *Journal of Econometrics* 31(3), pp. 307–327.

Bollerslev, T. (1987), A Conditionally Heteroskedastic Time Series Model for Speculative Prices and Rates of Return, *Review of Economics and Statistics* 69(3), pp. 542–547.

Bollerslev, T. (1988), On the Correlation Structure for the Generalized Autoregressive Conditional Heteroskedastic Process, *Journal of Time Series Analysis* 9(2), pp. 121–131.

Bollerslev, T. (2008), Glossary to ARCH (GARCH). CREATES Research Paper 2008-49, Center for Research in Econometric Analysis of Time Series, University of Aarhus.

Bollerslev, T., Chou, R. Y., and Kroner, K. F. (1992), ARCH Modeling in Finance, *Journal of Econometrics* 52(1), pp. 5–59.

Bollerslev, T., Engle, R. F., and Nelson, D. B. (1994), ARCH Models, in R. F. Engle and D. L. McFadden (eds.), *Handbook of Econometrics*, Amsterdam: Elsevier, vol. 4, pp. 2959–3038.

Bollerslev, T., Gibson, M., and Zhou, H. (2011), Dynamic Estimation of Volatility Risk Premia and Investor Risk Aversion from Option-Implied and Realized Volatilities, *Journal of Econometrics* 160(1), pp. 235–245.

Bouchaud, J.-P. and Potters, M. (2001), More Stylized Facts of Financial Markets: Leverage Effect and Downside Correlations, *Physica A: Statistical Mechanics and its Applications* 299(1-2), pp. 60–70.

Box, G. E. P., Jenkins, G. M., and Reinsel, G. C. (2008), *Time Series Analysis: Forecasting and Control*, Hoboken, NJ: John Wiley & Sons.

Boyle, P. P. and Vorst, T. (1992), Option Replication in Discrete Time with Transaction Costs, *The Journal of Finance* 47(1), pp. 271–293.

Brace, A. et al. (2002), Markovian Models in the Stochastic Implied Volatility Framework. Working Paper, University of New South Wales.

Branger, N. and Schlag, C. (2004), Why Is the Index Smile so Steep?, *Review of Finance* 8(1), pp. 109–127.

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

Brenner, M. and Subrahmanyan, M. G. (1988), A Simple Formula to Compute the Implied Standard Deviation, *Financial Analysts Journal* 44(5), pp. 80–83.

Brenner, R. (1996), Volatility is Not Constant, in I. Nelken (ed.), *The Handbook of Exotic Options: Instruments, Analysis, and Applications*, Chicago, IL: Irwin Professional, pp. 293–315.

Brigo, D. and Mercurio, F. (2001), *Interest Rate Models Theory and Practice*, Berlin: Springer.

Britten-Jones, M. and Neuberger, A. (2000), Option Prices, Implied Price Processes, and Stochastic Volatility, *The Journal of Finance* 55(2), pp. 839–866.

Brooks, C. (2008), *Introductory Econometrics for Finance*, Cambridge, UK: Cambridge UP.

Brooks, C. and Burke, S. P. (1998), Forecasting Exchange Rate Volatility Using Conditional Variance Models Selected by Information Criteria, *Economics Letters* 61(3), pp. 273–278.

Brooks, C. and Oozeer, C. M. (2002), Modelling the Implied Volatility of Options on Long Gilt Futures, *Journal of Business Finance & Accounting* 29(1-2), pp. 111–137.

Brooks, C. and Persand, G. (2003), Volatility Forecasting for Risk Management, *Journal of Forecasting* 22(1), pp. 1–22.

Brüggemann, R. et al. (2008), VAR Modeling for Dynamic Loadings Driving Volatility Strings, *Journal of Financial Econometrics* 6(3), pp. 361–381.

Brunner, B. and Hafner, R. (2003), Arbitrage-Free Estimation of the Risk-Neutral Density from the Implied Volatility Smile, *The Journal of Computational Finance* 7(1), pp. 75–106.

Buraschi, A. and Jiltsov, A. (2006), Model Uncertainty and Option Markets with Heterogeneous Beliefs, *The Journal of Finance* 61(6), pp. 2841–2897.

Busch, T., Christensen, B. J., and Nielsen, M. Ø. (2011), The Role of Implied Volatility in Forecasting Future Realized Volatility and Jumps in Foreign Exchange, Stock, and Bond Markets, *Journal of Econometrics* 160(1), pp. 48–57.

Campa, J. and Chang, P. H. K. (1995), Testing the Expectations Hypothesis on the Term Structure of Volatilities in Foreign Exchange Options, *The Journal of Finance* 50(2), pp. 529–547.

Campa, J. M., Chang, P. H. K., and Reider, R. L. (1997), ERM Bandwidths for EMU and After: Evidence from Foreign Exchange Options, *Economic Policy* 12(24), pp. 53–89.

Campa, J. M., Chang, P. H. K., and Reider, R. L. (1998), Implied Exchange Rate Distributions: Evidence from OTC Option Markets, *Journal of International Money and Finance* 17(1), pp. 117–160.

Campbell, J. Y., Lo, A. W., and MacKinlay, C. A. (1997), *The Econometrics of Financial Markets*, Princeton, NJ: Princeton UP.

Canina, L. and Figlewski, S. (1993), The Informational Content of Implied Volatility, *The Review of Financial Studies* 6(3), pp. 659–681.

Caporin, M. and McAleer, M. (2012), Model Selection and Testing of Conditional and Stochastic Volatility Models, in L. Bauwens, C. M. Hafner, and S. Laurent (eds.), *Handbook of Volatility Models and Their Applications*, Hoboken, NJ: John Wiley & Sons, vol. 3, pp. 199–222.

Cassese, G. and Guidolin, M. (2006), Modelling the Implied Volatility Surface: Does Market Efficiency matter? An Application to MIB30 Index Options, *International Review of Financial Analysis* 15(2), pp. 145–178.

Chan, W. H., Jha, R., and Kalimipalli, M. (2009), The Economic Value of Using Realized Volatility in Forecasting Future Implied Volatility, *The Journal of Financial Research* 32(3), pp. 231–259.

Chan, W.-S. (1995), Understanding the Effect of Time Series Outliers on Sample Autocorrelations, *Test* 4(1), pp. 179–186.

Cheng, X. and Fung, J. K. W. (2012), The Informational Content of Model-Free Implied Volatility, *The Journal of Futures Markets* 32(8), pp. 792–806.

Chernov, M. (2007), On the Role of Risk Premia in Volatility Forecasting, *The Journal of Business and Economic Studies* 25(4), pp. 411–426.

Chiras, D. P. and Manaster, S. (1978), The Information Content of Option Prices and a Test of Market Efficiency, *Journal of Financial Economics* 6(2-3), pp. 213–234.

Choi, K. and Zivot, E. (2007), Long Memory and Structural Changes in the Forward Discount: An Empirical Investigation, *Journal of International Money and Finance* 26(3), pp. 342–363.

Chow, G. C. (1960), Test of Equality between Sets of Coefficients in Two Linear Regressions, *Econometrica* 28(3), pp. 591–605.

Chriss, N. A. (1997), *Black-Scholes and Beyond: Option Pricing Models*, New York: McGraw-Hill.

Christensen, B. J. and Prabhala, N. R. (1998), The Relation between Implied and Realized Volatility, *Journal of Financial Economics* 50(2), pp. 125–150.

Christensen, K., Oomen, R. C. A., and Podolskij, M. (2014), Fact or Friction: Jumps at Ultra High Frequency, *Journal of Financial Economics* 114(3), pp. 576–599.

Christoffersen, P. and Jacobs, K. (2002), Which Volatility Model for Option Valuation? Working Paper, McGill Universtiy.

Claessen, H. and Mittnik, S. (2002), Forecasting Stock Market Volatility and the Informational Efficiency of the DAX-Index Options Market, *The European Journal of Finance* 8(3), pp. 302–321.

Clark, T. E. and McCracken, M. W. (2001), Tests for Equal Forecast Accuracy and Encompassing for Nested Models, *Journal of Econometrics* 105(1), pp. 85–110.

Clark, T. E. and McCracken, M. W. (2005), The Power of Tests of Predictive Ability in the Presence of Structural Breaks, *Journal of Econometrics* 124(1), pp. 1–31.

Coakley, J., Dollery, J., and Kellard, N. (2011), Long Memory and Structural Breaks in Commodity Futures Markets, *Journal of Futures Markets* 31(11), pp. 1076–1113.

Coleman, T. F., Li, Y., and Verma, A. (1999), Reconstructing the Unknown Local Volatility Function, *Journal of Computational Finance* 2(3), pp. 77–102.

Constantinides, G. M. (1996), Transaction Costs and the Implied Volatility Smile. Working Paper, Graduate School of Business, University of Chicago.

Constantinides, G. M., Jackwerth, J. C., and Perrakis, S. (2009), Mispricing of S&P 500 Index Options, *Review of Financial Studies* 22(3), pp. 1247–1277.

Cont, R. and da Fonseca, J. (2002), Dynamics of Implied Volatility Surfaces, *Quantitative Finance* 2(1), pp. 45–60.

Cont, R. and Tankov, P. (2004), *Financial Modelling with Jump Processes*, Boca Raton, FL: Chapman & Hall/CRC.

Corradi, V., Distaso, W., and Fernandes, M. (2012), International Market Links and Volatility Transmission, *Journal of Econometrics* 170(1), pp. 117–141.

Corradi, V. and Swanson, N. R. (2007), Nonparametric Bootstrap Procedures for Predictive Inference Based on Recursive Estimation Schemes, *International Economic Review* 48(1), pp. 67–109.

Corsi, F. (2009), A Simple Approximate Long-Memory Model of Realized Volatility, *Journal of Financial Econometrics* 13(4), pp. 174–196.

Corsi, F. et al. (2008), The Volatility of Realized Volatility, *Econometric Reviews* 27(1-3), pp. 46–78.

Cox, J. C. (1996), The Constant Elasticity of Variance Option Pricing Model, *The Journal of Portfolio Management* 23(1), pp. 15–17.

Cox, J. C., Ingersoll Jr, J. E., and Ross, S. A. (1985), An Intertemporal General Equilibrium Model of Asset Prices, *Econometrica* 53(2), pp. 363–384.

Cox, J. C. and Ross, S. A. (1976), The Valuation of Options for Alternative Stochastic Processes, *Journal of Financial Economics* 3(1-2), pp. 145–166.

Cox, J. C., Ross, S. A., and Rubinstein, M. (1979), Option Pricing: A Simplified Approach, *Journal of Financial Economics* 7(3), pp. 229–263.

Crato, N. and Ray, B. K. (196), Model Selection and Forecasting for Long-Range Dependent Processes, *Journal of Forecasting* 15(2), pp. 107–125.

Das, S. R. and Sundaram, R. K. (1999), Of Smiles and Smirks: A Term-Structure Perspective, *Journal of Financial Quantitative Analysis* 34(2), pp. 211–240.

Day, T. E. and Lewis, C. M. (1992), Stock Market Volatility and the Information Content of Stock Index Options, *Journal of Econometrics* 52(1-2), pp. 267–287.

Demeterfi, K. et al. (1999), More Than You Ever Wanted to Know about Volatility Swaps. Quantitative Strategies Research Notes, Goldman Sachs, New York.

Derman, E. and Kani, I. (1994a), Riding on a Smile, *Risk* 7(2), pp. 32–39.

Derman, E. and Kani, I. (1994b), The Volatility Smile and Its Implied Tree. Quantitative Strategies Research Notes, Goldman Sachs, New York.

Derman, E. and Kani, I. (1998), Stochastic Implied Trees: Arbitrage Pricing with Stochastic Term and Strike Structure of Volatility, *International Journal of Theoretical and Applied Finance* 1(1), pp. 61–110.

Derman, E., Kani, I., and Chriss, N. (1996), Implied Trinomial Trees of the Volatility Smile, *Journal of Derivatives* 3(4), pp. 7–22.

Detlefsen, K. (2007), *Equity Derivatives Markets*, Ph.D. thesis, Humboldt-Universität zu Berlin.

Detlefsen, K. and Härdle, W. K. (2007), Calibration Risk for Exotic Options, *The Journal of Derivatives* 14(4), pp. 47–63.

Diebold, F. X. (1988), *Empirical Modeling of Exchange Rate Dynamics*, vol. 303, Berlin: Springer.

Diebold, F. X. and Inoue, A. (2001), Long Memory and Regime Switching, *Journal of Econometrics* 105(1), pp. 131–159.

Diebold, F. X. and Mariano, R. S. (1995), Comparing Predictive Accuracy, *Journal of Business and Economic Statistics* 13(3), pp. 253–263.

Diz, F. and Finucane, T.J. (1993), Do the Options Markets Really Overreact?, *Journal of Futures Markets* 13(3), pp. 299–312.

Doornik, J. A. and Ooms, M. (2006), A Package for Estimating, Forecasting and Simulating ARFIMA Models: ARFIMA Package 1.04 for Ox. Working Paper, Nuffield College, Oxford.

Duan, J. C. (1999), Conditionally Fat-Tailed Distributions and the Volatility Smile in Options. Working Paper, Department of Finance, Hong Kong University of Science and Technology.

Duffie, D., Pan, J., and Singleton, K. (2000), Transform Analysis and Asset Pricing for Affine Jump-Diffusions, *Econometrica* 68(6), pp. 1343–1376.

Dumas, B., Fleming, J., and Whaley, R. E. (1998), Implied Volatility Functions: Empirical Tests, *The Journal of Finance* 53(6), pp. 2059–2106.

Dunis, C., Kellard, N. M., and Snaith, S. (2013), Forecasting EUR–USD Implied Volatility: The Case of Intraday Data, *Journal of Banking & Finance* 37(12), pp. 4943–4957.

Dupire, B. (1994), Pricing with a Smile, *RISK* 7(1), pp. 18–20.

Dupire, B. (1996), A Unified Theory of Volatility. Working Paper, Paribas Capital Markets.

Easley, D., O'Hara, M., and Srinivas, P. S. (1998), Option Volume and Stock Prices: Evidence on Where Informed Traders Trade, *The Journal of Finance* 53(2), pp. 431–466.

Ebens, H. (1999), Realized Stock Volatility. Working Paper 420, Department of Economics, Johns Hopkins University.

Ederington, L. and Guan, W. (2002), Is Implied Volatility an Informationally Efficient and Effective Predictor of Future Volatility?, *The Journal of Risk* 4(3), pp. 29–46.

Ederington, L. and Guan, W. (2005), The Information Frown in Option Prices, *Journal of Banking & Finance* 29(6), pp. 1429–1457.

Ekstrand, C. (2011), *Financial Derivatives Modeling*, Berlin: Springer.

Elliott, G., Rothenberg, T. J., and Stock, J. H. (1996), Efficient Tests for an Autoregressive Unit Root, *Econometrica* 64(4), pp. 813–836.

Ender, M. (2008), *Modellrisiko bei der Bewertung von Optionen in einem Vergleich von Modellen mit stochastischer Volatilität*, Ph.D. thesis, Universität zu Köln.

Enders, W. (2004), *Applied Econometric Time Series*, New York: John Wiley & Sons.

Engle, R. (1982), Autoregressive Condidtional Heteroskedascticity with Esimtates of the Variance of U.K. Inflation, *Econometrica* 50(4), pp. 987–1008.

Engle, R. F. and Bollerslev, T. (1986), Modeling the Persistence of Conditional Variances, *Econometric Reviews* 5(1), pp. 1–50.

Engle, R. F. and Ng, V. K. (1993), Measuring and Testing the Impact of News on Volatility, *The Journal of Finance* 48(5), pp. 1749–1778.

Engle, R. F. and Patton, A. J. (2001), What Good Is a Volatility Model, *Quantitative finance* 1(2), pp. 237–245.

Engle, R. F. and Sokalska, M. E. (2012), Forecasting Intraday Volatility in the US Equity Market. Multiplicative Component GARCH, *Journal of Financial Econometrics* 10(1), pp. 54–83.

EUREX (2011), Conditions for Trading at Eurex Deutschland and Eurex Zürich, `http://www.eurexchange.com/download/documents/regulations/trading_conditions/trading_conditions_en.pdf`.

EUREX (2012), Contract Specifications for Futures Contracts and Options Contracts at Eurex Deutschland and Eurex Zürich, `http://www.eurexchange.com/download/documents/regulations/contract_specifications/contract_specifications_en.pdf`.

EUREX (2013), EUREX Monthly Statistics December 2012, `http://www.eurexchange.com/blob/307292/0393c0092fc68719452ca0ebed8c9049/data/monthlystat_201212.pdf`.

Evnine, J. and Rudd, E. (1985), Index Options: The Early Evidence, *The Journal of Finance* 40(3), pp. 743–756.

Fan, S. M. (2010), *Multiple Tests für die Evaluation von Prognosemodellen*, *Reihe Quantitative Ökonomie*, vol. 164, Köln: Josef Eul.

Feinstein, S. (1989), *A Theoretical and Empirical Investigation of the Black–Scholes Implied Volatility*, Ph.D. thesis, Yale University, New Haven.

Fengler, M., Härdle, W., and Schmidt, P. (2002), Common Factors Governing VDAX Movements and the Maximum Loss, *Financial Markets and Portfolio Management* 16(1), pp. 16–29.

Fengler, M. R. (2004), *Semiparametric Modelling of Implied Volatility*, Berlin: Springer.

Fengler, M. R. (2009), Arbitrage-Free Smoothing of the Implied Volatility Surface, *Quantitative Finance* 9(4), pp. 417–428.

Fengler, M. R., Härdle, W. K., and Mammen, E. (2007), A Semiparametric Factor Model for Implied Volatility Surface Dynamics, *Journal of Financial Econometrics* 5(2), pp. 189–218.

Fengler, M. R., Härdle, W. K., and Villa, C. (2003), The Dynamics of Implied Volatilities: A Common Principal Component Approach, *Review of Derivatives Research* 6(3), pp. 179–202.

Fengler, W. (2012), Option Data and Modelling BSM Implied Volatility, in J.-C. Duan, W. K. Härdle, and J. E. Gentle (eds.), *Handbook of Computational Finance*, Berlin: Springer, pp. 117–142.

Figlewski, S. (1997), Forecasting Volatility, *Financial Markets, Institutions & Instruments* 6(1), pp. 1–88.

Fildes, R. and Makridakis, S. (1995), The Impact of Empirical Accuracy Studies on Time Series Analysis and Forecasting, *International Statistical Review* 63(3), pp. 289–308.

Fleming, J. (1998), The Quality of Market Volatility Forecasts Implied by S&P 100 Index Option Prices, *Journal of Empirical Finance* 5(4), pp. 317–345.

Fleming, J., Ostdiek, B., and Whaley, R. E. (1995), Predicting Stock Market Volatility: A New Measure, *The Journal of Futures Markets* 15(3), pp. 265–302.

Foresi, S. and Wu, L. (2005), Crash-O-Phobia: A Domestic Fear or a Worldwide Concern?, *Journal of Derivatives* 13(2), pp. 8–21.

Francq, C. and Zakoian, J.-M. (2010), *GARCH Models: Structure, Statistical Inference and Financial Applications*, Hoboken, NJ: John Wiley & Sons.

Franke, J., Härdle, W., and Hafner, C. M. (2004), *Statistics of Financial Markets: An Introduction*, Berlin: Springer.

Franke, J., Härdle, W. K., and Hafner, C. M. (2011), *Statistics of Financial Markets: An Introduction*, Berlin: Springer, 3rd ed.

Franks, J. R. and Schwartz, E. S. (1991), The Stochastic Behaviour of Market Variance Implied in the Prices of Index Options, *The Economic Journal* 101(409), pp. 1460–1475.

Frey, R. and Patie, P. (2002), Risk Management for Derivatives in Illiquid Markets: A Simulation Study, in K. Sandmann and P. J. Schönbucher (eds.), *Advances in Finance and Stochastics*, Berlin: Springer, pp. 137–159.

Fung, W. K. H. and Hsieh, D. A. (1991), Empirical Analysis of Implied Volatility: Stocks, Bonds, and Currencies. Working Paper, Duke University.

Garcia, R., Ghysels, E., and Renault, E. (2010), The Econometrics of Option Pricing, in Y. Aït-Sahalia and L. P. Hansen (eds.), *Handbook of Financial Econometrics*, Amsterdam: Elsevier, vol. 1, pp. 479–552.

Gatheral, J. (2006), *The Volatility Surface: A Practitioner's Guide*, Hoboken, NJ: John Wiley & Sons.

Geweke, J. and Porter-Hudak, S. (1983), The Estimation and Application of Long Memory Time Series Models, *Journal of Time Series Analysis* 4(4), pp. 221–238.

Gilster, J. E. and Lee, W. (1984), The Effects of Transaction Costs and Different Borrowing and Lending Rates on the Option Pricing Model: A Note, *The Journal of Finance* 39(4), pp. 1215–1221.

Glosten, L. R., Jagannathan, R., and Runkle, D. E. (1993), On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks, *The Journal of Finance* 48(5), pp. 1779–1801.

Gonçalves, S. and Guidolin, M. (2006), Predictable Dynamics in the S&P 500 Index Options Implied Volatility Surface, *The Journal of Business* 79(3), pp. 1591–1635.

Gospodinov, N. (2002), Median Unbiased Forecasts for Highly Persistent Autoregressive Processes, *Journal of Econometrics* 111(1), pp. 85–101.

Gospodinov, N., Gavala, A., and Jiang, D. (2006), Forecasting Volatility, *Journal of Forecasting* 25(6), pp. 381–400.

Gourieroux, C. and Jasiak, J. (2001a), *Financial Econometrics*, Princeton, NJ: Princeton UP.

Gourieroux, C. and Jasiak, J. (2001b), Memory and Infrequent Breaks, *Economics Letters* 70(1), pp. 29–41.

Granger, C. W. J. (1980), Long Memory Relationships and the Aggregation of Dynamic Models, *Journal of Econometrics* 14(2), pp. 227–238.

Granger, C. W. J. (1999), *Empirical Modeling in Economics: Specification and Evaluation*, Cambridge, UK: Cambridge UP.

Granger, C. W. J. and Ding, Z. (1996), Varieties of Long Memory Models, *Journal of Econometrics* 73(1), pp. 61–77.

Granger, C. W. J. and Hyung, N. (2004), Occasional Structural Breaks and Long Memory with an Application to the S&P 500 Absolute Stock Returns, *Journal of Empirical Finance* 11(3), pp. 399–421.

Granger, C. W. J. and Joyeux, R. (1980), An Introduction to Long-Memory Time Series Models and Fractional Differencing, *Journal of Time Series Analysis* 1(1), pp. 15–29.

Granger, C. W. J and Teräsvirta, T. (1999), A Simple Nonlinear Time Series Model with Misleading Linear Properties, *Economics Letters* 62(2), pp. 161–165.

Gregoriou, G. N. (2009), *Stock Market Volatility*, Boca Raton, FL: Chapman & Hall.

Grossman, S. J. and Zhou, Z. (1996), Equilibrium Analysis of Portfolio Insurance, *The Journal of Finance* 51(4), pp. 1379–1403.

Hafner, R. (2004), *Stochastic Implied Volatility*, Berlin: Springer.

Hafner, R. and Wallmeier, M. (2000), The Dynamics of DAX Implied Volatilities. Working Paper, University of Augsburg.

Hagan, P. S. et al. (2002), Managing Smile Risk, *Wilmott Magazine* 1, pp. 84–108.

Hansen, B. E. (1997), Approximate Asymptotic *P* Values for Structural Change Tests, *Journal of Business and Economic Statistics* 15(1), pp. 60–67.

Hansen, B. E. (2001), The New Econometrics of Structural Change: Dating Breaks in U.S. Labor Productivity, *Journal of Economic Perspectives* 15(4), pp. 117–128.

Hansen, B. E. (2012), Advanced Time Series and Forecasting, Structural Breaks. Summer School in Economics and Econometrics, July 23-27, 2012, University of Crete.

Hansen, P. R. (2005), A Test for Superior Predictive Ability, *Journal of Business and Economic Statistics* 23(4), pp. 365–380.

Hansen, P. R., Huang, Z., and Shek, H. H. (2011a), Realized GARCH: A Joint Model for Returns and Realized Measures of Volatility, *Journal of Applied Econometrics* 27(6), pp. 877–906.

Hansen, P. R. and Lunde, A. (2005), A Forecast Comparison of Volatility Models: Does Anything Beat a GARCH(1,1)?, *Journal of Applied Econometrics* 20(7), pp. 873–889.

Hansen, P. R. and Lunde, A. (2006), Consistent Ranking of Volatility Models, *Journal of Econometrics* 131(1-2), pp. 97–121.

Hansen, P. R. and Lunde, A. (2011), Forecasting Volatility Using High-Frequency Data, in M. P. Clements and D. F. Hendry (eds.), *The Oxford Handbook of Economic Forecasting*, New York: Oxford UP, pp. 525–558.

Hansen, P. R., Lunde, A., and Nason, J. M. (2003), Choosing the Best Volatility Models: The Model Confidence Set Approach, *Oxford Bulletin of Economics and Statistics* 65(s1), pp. 839–861.

Hansen, P. R., Lunde, A., and Nason, J. M. (2011b), The Model Confidence Set, *Econometrica* 79(2), pp. 453–497.

Härdle, W. (1990), *Applied Nonparametric Regression*, Cambridge, UK: Cambridge UP.

Härdle, W., Hautsch, N., and Pigorsch, U. (2008), Measuring and Modeling Risk Using High-Frequency Data, in W. Härdle, N. Hautsch, and L. Overbeck (eds.), *Applied Quantitative Finance*, Berlin: Springer, pp. 275–294.

Härdle, W., Kleinow, T., and Stahl, G. (2002), *Applied Quantitative Finance: Theory and Computational Tools*, Berlin: Springer.

Harvey, C. R. and Whaley, R. E. (1991), S&P 100 Index Option Volatility, *The Journal of Finance* 46(4), pp. 1551–1561.

Harvey, C. R. and Whaley, R. E. (1992), Market Volatility Prediction and the Efficiency of the S&P 100 Index Option Market, *Journal of Financial Economics* 31(1), pp. 43–73.

Harvey, D. (1997), *The Evaluation of Economic Forecasts*, Ph.D. thesis, University of Nottingham.

Harvey, D., Leybourne, S., and Newbold, P. (1997), Testing the Equality of Prediction Mean Squared Errors, *International Journal of Forecasting* 13(2), pp. 281–291.

Harvey, D. and Newbold, P. (2000), Tests for Multiple Forecast Encompassing, *Journal of Applied Econometrics* 15(5), pp. 471–482.

Haug, E. G. (2007), *The Complete Guide to Option Pricing Formulas*, New York: McGraw-Hill.

Hautsch, N. (2012), *Econometrics of Financial High-Frequency Data*, Berlin: Springer.

Herrmann, R. (1999), *Nichtparametrische Optionsbewertung*, Fankfurt am Main: Peter Lang.

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

Heynen, R., Kemna, A., and Vorst, T. (1994), Analysis of the Term Structure of Implied Volatilities, *Journal of Financial and Quantitative Analysis* 29(1), pp. 31–56.

Hidalgo, J. and Robinson, P. M. (1996), Testing for Structural Change in a Long-Memory Environment, *Journal of Econometrics* 70(1), pp. 159–174.

Horn, A. (2012), *Dynamik und Glättung von impliziten Volatilitätsoberflächen*, Master's thesis, Philipps-Universität Marburg.

Hosking, J. R. M. (1981), Fractional Differencing, *Biometrika* 68(1), pp. 165–176.

Hsieh, D. A. (1989), Testing for Nonlinear Dependence in Daily Foreign Exchange Rates, *The Journal of Business* 62(3), pp. 339–368.

Hull, J. C. (2006), *Options, Futures, and Other Derivatives*, Upper Saddle River, NJ: Pearson Prentice Hall.

Hull, J. C. and White, A. (1987), The Pricing of Options on Assets with Stochastic Volatilities, *The Journal of Finance* 42(2), pp. 281–300.

Hurvich, C. M., Simonoff, J. S., and Tsai, C.-L. (1998), Smoothing Parameter Selection in Nonparametric Regression Using an Improved Akaike Information Criterion, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 60(2), pp. 271–293.

Inoue, A. and Kilian, L. (2005), In-Sample or Out-of-Sample Tests of Predictability: Which One Should We Use?, *Econometric Reviews* 23(4), pp. 371–402.

Isaenko, S. (2007), Dynamic Equilibrium with Overpriced Put Options, *Economic Notes* 36(1), pp. 1–26.

Jackwerth, J. C. (2004), *Option-Implied Risk-Neutral Distributions and Risk Aversion*, Charlottesville, VA: The Research Foundation of AIMR.

Jackwerth, J. C. and Rubinstein, M. (2001), Recovering Stochastic Processes from Option Prices. Working Paper, London Business Schoool.

Jarrow, R. A. (1998), *Volatility: New Estimation Techniques for Pricing Derivatives*, London: Risk Books.

Javaheri, A. (2005), *Inside Volatility Arbitrage*, Hoboken, NJ: John Wiley & Sons.

Javaheri, A., Wilmott, P., and Haug, E. G. (2004), GARCH and Volatility SWAPs, *Quantitative Finance* 4(5), pp. 589–595.

Jiang, G. J. and Tian, Y. S. (2005), The Model-Free Implied Volatility and Its Information Content, *Review of Financial Studies* 18(4), pp. 1305–1342.

Jiang, G. J. and Tian, Y. S. (2010), Misreaction or Misspecification? A Re-Examination of Volatility Anomalies, *Journal of Banking & Finance* 34(10), pp. 2358–2369.

Joshi, M. S. (2003), *The Concepts and Practice of Mathematical Finance*, Cambridge, UK: Cambridge UP.

Kennedy, P. (2003), *A Guide to Econometrics*, Cambridge, MA: MIT.

Kim, C. S. and Phillips, P. C. B. (2006), Log Periodogram Regression: The Nonstationary Case. Cowles Foundation Discussion Papers 1587, Cowles Foundation for Research in Economics, Yale University.

Kirchgässner, G. and Wolters, J. (2007), *Introduction to Modern Time Series Analysis*, Berlin: Springer.

Klein, L. R. (1992), The Test of a Model Is Its Ability to Predict. Working Paper, University of Pennsylvania.

Kokoszka, P. S. and Politis, D. N. (2008), The Variance of Sample Autocorrelations: Does Barlett's Formula Work with ARCH Data? Working Paper, Utah State University and University of California at San Diego.

Konstantinidi, E., Skiadopoulos, G., and Tzagkaraki, E. (2008), Can the Evolution of Implied Volatility Be Forecasted? Evidence from European and US Implied Volatility Indices, *Journal of Banking & Finance* 32(11), pp. 2401–2411.

Koopman, S. J., Jungbacker, B., and Hol, E. (2005), Forecasting Daily Variability of the S&P 100 Stock Index Using Historical, Realised and Implied Volatility Measurements, *Journal of Empirical Finance* 12(3), pp. 445–475.

Krämer, W. and Runde, R. (1994), Some Pitfalls in Using Empirical Autocorrelations to Test for Zero Correlation Among Common Stock Returns, in J. Kaehler and P. Kugler (eds.), *Econometric Analysis of Financial Markets*, Heidelberg: Physica, pp. 1–10.

Kuan, C.-M. and Hsu, C.-C. (1998), Change-Point Estimation of Fractionally Integrated Processes, *Journal of Time Series Analysis* 19(6), pp. 693–708.

Kwiatkowski, D. et al. (1992), Testing the Null Hypothesis of Stationarity Against the Alternative of a Unit Root, *Journal of Econometrics* 54(1-3), p. 159–178.

Lamoureux, C. G. and Lastrapes, W. D. (1990), Persistence in Variance, Structural Change, and the GARCH Model, *Journal of Business and Economic Statistics* 8(2), pp. 225–234.

Lamoureux, C. G. and Lastrapes, W. D. (1993), Forecasting Stock-Return Variance: Toward an Understanding of Stochastic Implied Volatilities, *The Review of Financial Studies* 6(2), pp. 293–326.

Landler, M. and Timmons, H. (2008), Stocks Plunge Worldwide on Fears of a U.S. Recession, *The New York Times*, January 21, 2008, `http://www.nytimes.com/2008/01/21/business/22stox-web.html?pagewanted=all&_r=0`.

Latané, H. A. and Rendleman, R. J. (1976), Standard Deviations of Stock Price Ratios Implied in Option Prices, *The Journal of Finance* 31(2), pp. 369–381.

Lazarov, Z. (2004), *Three Essays on the Econometrics of Options Markets*, Berlin: dissertation.de.

Ledoit, O., Santa-Clara, P., and Yan, S. (2002), Relative Pricing of Options with Stochastic Volatility. Working Paper, Anderson Graduate School of Management, University of California at Los Angeles.

Leeb, H. and Pötscher, B. M. (2009), Model Selection, in T. Mikosch, J.-P. Kreiß, R. A. David, and T. G. Andersen (eds.), *Handbook of Financial Time Series*, Berlin: Springer, pp. 889–925.

Leland, H. E. (1985), Option Pricing and Replication with Transactions Costs, *The Journal of Finance* 40(5), pp. 1283–1301.

Li, K. (2002), Long-Memory versus Option-Implied Volatility Predictions, *Journal of Derivatives* 9(3), pp. 9–25.

Liu, L., Patton, A. J., and Sheppard, K. (2012), Does Anything Beat 5-Minute RV? Discussion Paper 645, Department of Economics, University of Oxford.

Ljung, G. M. and Box, G. E. P. (1978), On a Measure of Lack of Fit in Time Series Models, *Biometrika* 65(2), pp. 297–303.

Longstaff, F. A. (1995), Option Pricing and the Martingale Restriction, *The Review of Financial Studies* 8(4), pp. 1091–1124.

Longstaff, F. A. (2005), Asset Pricing in Markets with Illiquid Assets. Working Paper, University of California.

Louzis, D. P., Xanthopoulos-Sisinis, S., and Refenes, A. P. (2012), Stock Index Realized Volatility Forecasting in the Presence of Heterogeneous Leverage Effects and Long Range Dependence in the Volatility of Realized Volatility, *Applied Economics* 44(27), pp. 3533–3550.

Louzis, D. P., Xanthopoulos-Sisinis, S., and Refenes, A. P. (2014), Realized Volatility Models and Alternative Value-at-Risk Prediction Strategies, *Economic Modelling* 40, pp. 101–116.

MacBeth, J. and Merville, L. (1979), An Empirical Examination of the Black-Scholes Call Option Pricing Model, *The Journal of Finance* 34(5), pp. 1173–1186.

Maddala, G. S. (2008), *Unit Roots, Cointegration, and Structural Change*, Cambridge, UK: Cambridge UP.

Makridakis, S. and Hibon, M. (2000), The M3-Competition: Results, Conclusions and Implications, *International Journal of Forecasting* 16(4), pp. 451–476.

Malkiel, B. (1992), Efficient Market Hypothesis, in P. Newman, M. Milgate, and J. Eatwell (eds.), *New Palgrave Dictionary of Money and Finance*, London: Palgrave Macmillan, pp. 739–744.

Marcellino, M. (2004), Forecast Pooling for European Macroeconomic Variables, *Oxford Bulletin of Economics and Statistics* 66(1), pp. 91–112.

Mariano, R. S. (2004), Testing Forecast Accuracy, in M. P. Clements and D. F. Hendry (eds.), *A Companion to Economic Forecasting*, Malden, MA: Blackwell, pp. 284–298.

Martens, M., Van Dijk, D., and De Pooter, M. (2009), Forecasting S&P 500 Volatility: Long Memory, Level Shifts, Leverage Effects, Day-of-the-Week Seasonality, and Macroeconomic Announcements, *International Journal of Forecasting* 25(2), pp. 282–303.

Martens, M. and Zein, J. (2004), Predicting Financial Volatility: High-Frequency Time-Series Forecasts Vis-à-Vis Implied Volatility, *The Journal of Futures Markets* 24(11), pp. 1005–1028.

Martin, G. M., Reidy, A., and Wright, J. (2009), Does the Option Market Produce Superior Forecasts of Noise-Corrected Volatility Measures?, *Journal of Applied Econometrics* 24(1), pp. 77–104.

Martino, S. et al. (2011), Estimating Stochastic Volatility Models Using Integrated Nested Laplace Approximations, *The European Journal of Finance* 17(7), pp. 487–503.

Mayhew, S. (1995), Implied Volatility, *Financial Analysts Journal* 51(4), pp. 8–20.

McAleer, M., Jimenez-Martin, J.-A., and Perez-Amaral, T. (2013), GFC-Robust Risk Management Strategies Under the Basel Accord, *International Review of Economics & Finance* 27, pp. 97–111.

McAleer, M. and Medeiros, M. C. (2008), Realized Volatility: A Review, *Econometric Reviews* 27(1-3), pp. 10–45.

McCracken, M. W. and West, K. D. (2004), Inference about Predictive Ability, in M. P. Clements and D. F. Hendry (eds.), *A Companion to Economic Forecasting*, Malden, MA: Blackwell, pp. 299–321.

Merton, R. (1973), The Theory of Rational Option Pricing, *Bell Journal of Economics and Management Science* 4(1), pp. 141–183.

Merton, R. C. (1976), Option Pricing When Underlying Stock Returns are Discontinuous, *Journal of Financial Economics* 3(1-2), pp. 125–144.

Merton, R. C. (1980), On Estimating the Expected Return on the Market: An Exploratory Investigation, *Journal of Financial Economics* 8(4), pp. 323–361.

Mitra, S. (2009), A Review of Volatility and Option Pricing. Working Paper, Brunel University.

Mittnik, S. and Rieken, S. (2000), Put-Call Parity and the Informational Efficiency of the German DAX-Index Options Market, *International Review of Financial Analysis* 9(3), pp. 259–279.

Mixon, S. (2007), The Implied Term Structure of Stock Index Options, *Journal of Empricial Finance* 14(3), pp. 333–354.

Mixon, S. (2009), Option Markets and Implied Volatility: Past versus Present, *Journal of Financial Economics* 94(2), pp. 171–191.

Morgan, J. P. (1996), *Riskmetrics: Technical Document*, New York: Morgan Guaranty Trust Company of New York.

Müller, U. A. et al. (1995), Fractals and Intrinsic Time: A Challenge to Econometricians. Working Paper, XXXIXth International AEA Conference on Real Time Econometrics, 14-15 October 1993, Reprint by Olsen and Associates, Zürich.

Müller, U. A. et al. (1997), Volatilities of Different Time Resolutions-Analyzing the Dynamics of Market Components, *Journal of Empirical Finance* 4(2-3), pp. 213–239.

Musiela, M. and Rutkowski, M. (2005), *Martingale Methods in Financial Modelling*, Berlin: Springer.

Muzzioli, S. (2010), Option Based Forecasts of Volatility: An Empirical Study in the DAX-Index Options Market, *The European Journal of Finance* 16(6), pp. 561–586.

Nadaraya, E. A. (1964), On Estimating Regression, *Theory of Probability & Its Applications* 9(1), pp. 141–142.

Nagel, H. (2001), *Optionsbewertung bei stochastischer Volatilität*, Wiesbaden: Deutscher Universitäts-Verlag.

Nandi, S. (2000), Asymmetric Information about Volatility: How Does it Affect Implied Volatility, Option Prices and Market Liquidity?, *Review of Derivatives Research* 3(3), pp. 215–236.

Ncube, M. (1996), Modelling Implied Volatility with OLS and Panel Data Models, *Journal of Banking & Finance* 20(1), pp. 71–84.

Nelson, D. B. (1991), Conditional Heteroskedasticity in Asset Returns: A New Approach, *Econometrica* 59(2), pp. 347–370.

Neumann, M. (1999), *Optionsbewertung und Risikomessung mit impliziten Binomialbäumen*, Wiesbaden: Deutscher Universitäts-Verlag.

Pagan, A. R. and Schwert, G. W. (1990), Alternative Models for Conditional Stock Volatility, *Journal of Econometrics* 45(1-2), pp. 267–290.

Pan, J. (2002), The Jump-Risk Premia Implicit in Options: Evidence From an Integrated Time-Series Study, *Journal of Financial Economics* 63(1), pp. 3–50.

Parkinson, M. (1980), The Extreme Value Method for Estimating the Variance of the Rate of Return, *The Journal of Business* 53(1), pp. 61–65.

Patton, A. J. (2011), Volatility Forecast Comparison Using Imperfect Volatility Proxies, *Journal of Econometrics* 160(1), pp. 246–256.

Patton, A. J. and Sheppard, K. (2009), Optimal Combinations of Realised Volatility Estimators, *International Journal of Forecasting* 25(2), pp. 218–238.

Pena, I., Rubio, G., and Serna, G. (1999), Why Do We Smile? On the Determinants of the Implied Volatility Function, *Journal of Banking & Finance* 23(8), pp. 1151–1179.

Pesaran, M. H. and Timmermann, A. (2007), Selection of Estimation Window in the Presence of Breaks, *Journal of Econometrics* 137(1), pp. 134–161.

Platen, E. and Schweizer, M. (1998), On Feedback Effects from Hedging Derivatives, *Mathematical Finance* 8(1), pp. 67–84.

Politis, D. N. and Romano, J. P. (1994), The Stationary Bootstrap, *Journal of the American Statistical Association* 89(428), pp. 1303–1313.

Pong, S. et al. (2004), Forecasting Currency Volatility: A Comparison of Implied Volatilities and AR(FI)MA Models, *Journal of Banking & Finance* 28(10), pp. 2541–2563.

Poon, S.-H. (2005), *A Practical Guide to Forecasting Financial Market Volatility*, Chichester, UK: John Wiley & Sons.

Poon, S.-H. and Granger, C. (2003), Forecasting Volatility in Financial Markets: A Review, *Journal of Economic Literature* 41(2), pp. 478–539.

Poon, S.-H. and Granger, C. (2005), Practical Issues in Forecasting Volatility, *Financial Analysts Journal* 61(1), pp. 45–56.

Poterba, J. M. and Summers, L. H. (1986), The Persistence of Volatility and Stock Market Fluctuations, *American Economic Review* 76(5), pp. 1142–1151.

Poteshman, A. M. (2000), Forecasting Future Volatility from Option Prices. Working Paper, University of Illinois at Urbana-Champaign.

Quandt, R. (1960), Tests of the Hypothesis that a Linear Regression System Obeys Two Separate Regimes, *Journal of the American Statistical Association* 55(290), pp. 324–330.

Raunig, B. (2006), The Longer-Horizon Predictability of German Stock Market Volatility, *International Journal of Forecasting* 22(2), pp. 363–372.

Rebonato, R. (2004), *Volatility and Correlation: The Perfect Hedger and the Fox*, Chichester, UK: John Wiley & Sons.

Renault, E. and Touzi, N. (1996), Option Hedging and Implied Volatilities in a Stochastic Volatility Model, *Mathematical Finance* 6(3), pp. 279–302.

Ripper, K. and Günzel, A. (1997), Volatilitäts-Smile von DAX-Optionen, *Finanzmarkt und Portfolio Management* 11(4), pp. 470–477.

Robinson, P. M. (1995a), Gaussian Semiparametric Estimation of Long Range Dependence, *The Annals of Statistics* 23(5), pp. 1630–1661.

Robinson, P. M. (1995b), Log-Periodogram Regression of Time Series with Long Range Dependence, *The Annals of Statistics* 23(3), pp. 1048–1072.

Rogers, L. C. G. and Tehranchi, M. R. (2010), Can the Implied Volatility Surface Move by Parallel Shifts?, *Finance and Stochastics* 14(2), pp. 235–248.

Roll, R. (1984), A Simple Implicit Measure of the Effective Bid-Ask Spread in an Efficient Market, *The Journal of Finance* 39(4), pp. 1127–1139.

Rosenberg, J. V. (1999), Implied Volatility Functions: A Reprise. Working Paper, University of New York.

Rouah, F. D. and Vainberg, G. (2007), *Option Pricing Models and Volatility Using Excel-VBA*, Hoboken, NJ: John Wiley & Sons.

Rubinstein, M. (1985), Non-Parametric Tests of Alternative Option Pricing Models Using all Reported Trades and Quotes on the 30 Most Active CBOE Option Classes from August 23, 1976 through August 31, 1978, *The Journal of Finance* 40(2), pp. 455–480.

Rubinstein, M. (1994), Implied Binomial Trees, *The Journal of Finance* 49(3), pp. 771–818.

Ruppert, D. and Wand, M. P. (1994), Multivariate Locally Weighted Least Squares Regression, *The Annals of Statistics* 22(3), pp. 1346–1370.

Sapusek, A. (2004), Volatility Clustering in German Stock Returns, in D. Gramlich and R. Schmidt (eds.), *Kapitalmarktforschung und Bankmanagement Impulse für eine anwendungsorientierte Forschung*, Wiesbaden: Deutscher Universitätsverlag, pp. 81–93.

Saunders, D. (1997), *Applications of Optimization to Mathematical Finance*, Ph.D. thesis, University of Toronto.

Schmalensee, R. and Trippi, R. R. (1978), Common Stock Volatility Expectations Implied by Option Premia, *The Journal of Finance* 33(1), pp. 129–147.

Schnellen, M. (2007), *Analysis of Implied Volatility Surfaces*, Ph.D. thesis, Universität Göttingen.

Schönbucher, P. J. (1999), A Market Model for Stochastic Implied Volatility, *Philosophical Transactions of the Royal Society A* 357(1758), pp. 2071–2092.

Schöne, A. (2010), Zum Informationsgehalt der Volatilitätsindizes VDAX und VDAX-New der Deutsche Börse AG, *Zeitschrift für betriebeswirtschaftliche Forschung* 62(Sep. 2010), pp. 625–661.

Schwert, G. W. (1989), Tests for Unit Roots: A Monte Carlo Investigation, *Journal of Business and Economic Statistics* 20(1), pp. 5–17.

Schwert, G. W. (1990), Stock Volatility and the Crash of '87, *The Review of Financial Studies* 3(1), pp. 77–102.

Schwert, G. W. (2011), Stock Volatility During the Recent Financial Crisis, *European Financial Management* 17(5), pp. 789–805.

Scott, L. O. (1997), Pricing Stock Options in a Jump-Diffusion Model with Stochastic Volatility and Interest Rates: Applications of Fourier Inversion Methods, *Mathematical Finance* 7(4), pp. 413–426.

Shefrin, H. (1999), Irrational Exuberance and Option Smiles, *Financial Analysts Journal* 55(6), pp. 91–103.

Shimko, D. C. (1993), Bounds on Probability, *Risk* 6(4), pp. 33–37.

Shu, J. and Zhang, J. E. (2003), The Relationship between Implied and Realized Volatility of S&P 500 Index, *Wilmott Magazine* 4(Jan.), pp. 83–91.

Skiadopoulos, G. (2001), Volatility Smile Consistent Option Models: A Survey, *International Journal of Theoretical and Applied Finance* 4(3), pp. 403–437.

Stein, J. (1989), Overreactions in the Options Market, *The Journal of Finance* 44(4), pp. 1011–1023.

Stock, J. H. and Watson, M. W. (1999), A Comparison of Linear and Nonlinear Univariate Models for Forecasting Macroeconomic Time Series, in R. F. Engle and H. White (eds.), *Cointegration, Causality, and Forecasting: A Festschrift in Honour of Clive W. J. Granger*, New York: Oxford UP, pp. 1–44.

Szakmary, A. et al. (2003), The Predictive Power of Implied Volatility: Evidence from 35 Futures Markets, *Journal of Banking & Finance* 27(11), pp. 2151–2175.

Tallau, C. (2011), Volatilitätsprognosen auf Basis der DAX-Volatilitätsindizes, *Kredit und Kapital* 44(1), pp. 47–74.

Taylor, S. J., Yadav, P. K., and Zhang, Y. (2010), The Information Content of Implied Volatilities and Model-Free Volatility Expectations: Evidence from Options Written on Indiviudal Stocks, *Journal of Banking & Finance* 34(4), pp. 871–881.

Teräsvirta, T. (2009), An Introduction to Univariate GARCH Models, in T. Mikosch, J.-P. Kreiß, R. A. David, and T. G. Andersen (eds.), *Handbook of Financial Time Series*, Berlin: Springer, pp. 17–42.

Timmermann, A. (2006), Forecast Combinations, in G. Elliott, C. W. J. Granger, and A. Timmermann (eds.), *Handbook of Economic Forecasting*, Amsterdam: Elsevier, vol. 1, pp. 135–196.

Tompkins, R. G. (2001), Implied Volatility Surfaces: Uncovering Regularities for Options on Finanical Futures, *The European Journal of Finance* 7(3), pp. 198–230.

Trautmann, S. (1990), Aktienoptionspreise an der Frankfurter Optionsbörse im Lichte der Optionsbewertungstheorie, in H. Göppl, W. Bühler, and R. von Rosen (eds.), *Optionen und Futures*, Frankfurt am Main: Knapp, pp. 79–100.

Trautmann, S. and Beinert, M. (1995), Stock Price Jumps and Their Impact on Option Valuation. Working Paper, Johannes-Gutenberg-Universität Mainz.

Tsay, R. S. (2005), *Analysis of Financial Time Series*, New York: John Wiley & Sons.

Vanden, J. M. (2008), Information Quality and Options, *Review of Financial Studies* 21(6), pp. 2635–2676.

Varian, H. (1975), A Bayesian Approach to Real Estate Assessment, in S. E. Feinberg and A. Zellner (eds.), *Studies in Bayesian Econometrics and Statistics in Honor of L. J. Savage*, Amsterdam: North Holland, pp. 195–208.

Viscusi, G. and Chassany, A.-S. (2008), Societe Generale Reports EU4.9 Billion Trading Loss, *Bloomberg*, January 24, 2008, `http://www.bloomberg.com/apps/news?pid=newsarchive&sid=a8GBEB7UuuXc&refer=home`.

Wallmeier, M. (2003), *Der Informationsgehalt von Optionspreisen*, Heidelberg: Physica.

Walter, B. (2008), *The Equity Volatility Smile and Default Risk*, Ph.D. thesis, Technische Universität München.

Wang, C. S.-H., Bauwens, L., and Hsiao, C. (2013), Forecasting a Long Memory Process Subject to Structural Breaks, *Journal of Econometrics* 177(2), pp. 171–184.

Watson, G. S. (1964), Smooth Regression Analysis, *Sankhyā: The Indian Journal of Statistics, Series A* 26(4), pp. 359–372.

West, K. D. (1996), Asymptotic Inference about Predictive Ability, *Econometrica* 64(5), pp. 1067–1084.

West, K. D. (2006), Forecast Evaluation, in G. Elliott, C. W. J. Granger, and A. Timmermann (eds.), *Handbook of Economic Forecasting*, Amsterdam: Elsevier, vol. 1, pp. 99–134.

West, K. D. and Cho, D. (1995), The Predictive Ability of Several Models of Exchange Rate Volatility, *Journal of Econometrics* 69(2), pp. 367–391.

Whaley, R. E. (2003), Derivatives, in G. M. Constantinides, M. Harris, and R. M. Stulz (eds.), *Handbook of the Economics of Finance*, Amsterdam: Elsevier, vol. 1, Part B, pp. 1129–1206.

White, H. (2000), A Reality Check for Data Snooping, *Econometrica* 68(5), pp. 1097–1126.

Wilhelmsson, A. (2006), Garch Forecasting Performance under Different Distribution Assumptions, *Statistical Methodology* 25(8), pp. 561–578.

Wilkens, S. and Röder, K. (2006), The Informational Content of Option-Implied Distributions: Evidence from the EUREX Index and Interest Rate Futures Options Market, *Global Finance Journal* 17(1), pp. 50–74.

Wilmott, P. (1998), *Derivatives: The Theory and Practice of Financial Engineering*, Chichester, UK: John Wiley & Sons.

Wilmott, P., Dewynne, J., and Howison, S. (1993), *Option Pricing: Mathematical Models and Computation*, Oxford, UK: Oxford Financial.

Xiao, L. (2013), Realized Volatility Forecasting: Empirical Evidence from Stock Market Indices and Exchange Rates, *Applied Financial Economics* 23(1), pp. 57–69.

Xiao, L. and Aydemir, A. (2011), Volatility Modelling and Forecsting in Finance, in S. Satchell and J. Knight (eds.), *Forecasting Volatility in the Financial Markets*, Oxford, UK: Butterworth-Heinemann, pp. 1–46, 3rd ed.

Xu, X. and Taylor, S. J. (1994), The Term Structure of Volatility Implied by Foreign Exchange Options, *Journal of Financial and Quantitative Analysis* 29(1), pp. 57–74.

Xu, X. and Taylor, S. J. (1995), Conditional Volatility and the Informational Efficiency of the PHLX Currency Options Market, *Journal of Banking & Finance* 19(5), pp. 803–821.

Yu, W. W., Lui, E. C. K., and Wang, J. W. (2010), The Predictive Power of the Implied Volatility of Options Traded OTC and on Exchanges, *Journal of Banking & Finance* 34(1), pp. 1–11.

Zakoian, J.-M. (1994), Threshold Heteroskedastic Models, *Journal of Economic Dynamics and Control* 18(5), pp. 931–955.

Zeileis, A. et al. (2003), Testing and Dating of Structural Changes in Practice, *Computational Statistics & Data Analysis* 44(1-2), pp. 109–123.

Zhang, L., Mykland, P. A., and Aït-Sahalia, Y. (2005), A Tale of Two Time Scales: Determining Integrated Volatility with Noisy High-Frequency Data, *Journal of the American Statistical Association* 100(472), pp. 1394–1411.

Zhu, J. (2000), *Modular Pricing of Options: An Application of Fourier Analysis*, Lecture Notes in Economics and Mathematical Systems, Berlin: Springer.

Zhu, Y. and Avellaneda, M. (1997), An E-ARCH Model for the Term Structure of Implied Volatility of FX Options, *Applied Mathematical Finance* 4(2), pp. 81–100.

Zivot, E. (2009), Practical Issues in the Analysis of Univariate GARCH Models, in T. Mikosch, J.-P. Kreiß, R. A. David, and T. G. Andersen (eds.), *Handbook of Financial Time Series*, Berlin: Springer, pp. 113–155.

Zivot, E. and Wang, J. (2008), *Modeling Financial Time Series with S-PLUS*, New York: Springer.