An Empirical Analysis of the DAX Index Options Market Using the GARCH Option Valuation Model of Heston & Nandi (2000)

Andreas Kanz*

Abgabedatum: 30.09.2016

Masterarbeit

am Lehrstuhl für BWL, insbes. Banking und Finanzcontrolling
der Otto-Friedrich-Universität Bamberg
unter Betreuung von Prof. Dr. Matthias Muck

[redacted]

Contents

List of Figures

List of Tables

List of Abbreviations

ARCH	Autoregressive Conditional Heteroskedasticity
ATM	At-the-Money
BSM	Black, Scholes and Merton
CF	Characteristic Function
CHJ	Christoffersen, Heston and Jacobs
CIR	Cox, Ingersoll and Ross
DAX	Deutscher Aktienindex
EGARCH	Exponential GARCH
GARCH	Generalized Autoregressive Conditional Heteroskedasticity
GJR-GARCH	Glosten, Jagannathan and Runkle GARCH
HN	Heston and Nandi
HNGARCH	Heston and Nandi GARCH
IGARCH	Integrated GARCH
ITM	In-the-Money
IVRMSE	Implied Volatility Root Mean Squared Error
MLE	Maximum Likelihood Estimation
NGARCH	Non-Linear Asymmetric GARCH
NIC	News Impact Curve
OTM	Out-of-the-Money
RMSE	Root Mean Squared Error
S&P500	Standard & Poor's 500
TTM	Time to Maturity

List of Symbols

α, β, γ, μ, ω	Parameters of the Heston & Nandi (2000) Model
C_t	Value of a Call Option at Time t
Δ	Number of Contracts in a Delta Hedge
δ, η, ϕ, $\hat{s}$, ϑ, ζ	Auxiliary Parameters
$\mathbb{E}$	Expectation
ε	Residual Term
h_t	Conditional Variance at Time t
$\mathcal{H}$	Half-Life of a Return Shock
I	Fisher Information Matrix
ι	Imaginary Unit
K	Strike Price of an Option
κ, ρ, θ, ϱ	Time-Invariant Parameters of the Heston (1993) Model
λ	Market Price of Volatility Risk in the Heston (1993) Model
$\mathcal{L}$	Likelihood Function
M	Stochastic Discount Factor
$\mathcal{N}$	Normal Distribution
∇	Length of a Time Step
Ω	Annualized Long-Run Volatility
p	Number of Autoregressive Terms in a GARCH Model
P_t	Value of a Put Option at Time t
$\mathbb{P}$	Empirical Probability Measure
Φ	Cumulative Standard Normal Distribution Function
Π	Probability Used to Obtain Option Prices
Ψ	Value of the Replicating Portfolio
q	Number of Moving Average Terms in a GARCH Model
$\mathbb{Q}$	Risk-Neutral Probability Measure
r	Risk-Free Interest Rate
R_t	Return from $t - \nabla$ until t
$\Re e$	Real Part of an Expression
S_t	Price of the Underlying at Time t
σ	Black-Scholes-Merton Model Implied Volatility
t	Time
T	Time at Expiration
τ	Time Period between t and T

Θ	Parameter Vector
υ	Drift of the Underlying
ν	State Variable for the Initial Variance in the Heston (1993) Model
W	Wiener Process
x	Logarithm of the Spot Price
ξ	Independent Variance Risk Premium Parameter
z	Standard Normal Innovation

1 Introduction

Originating from the works of Black and Scholes (1973) and Merton (1973), large numbers of different approaches have been suggested since, in attempts to price financial options. Tackling a well-known weakness of their seminal model, which is the constant volatility assumption, most of the attention has been given to models introducing stochastic volatility in a continuous-time framework. This concept turned out to be a major improvement and with advancing refinements through works of, for instance, Hull and White (1987), Stein and Stein (1991) or Heston (1993) stochastic volatility models have since been able to largely explain the skew in implied volatilities across moneyness. In addition to the stochastic volatility approach, another class of models operating in discrete-time has been used to price options. Following Amin and Ng (1993) and Duan (1995), Generalized Autoregressive Conditional Heteroskedasticity (GARCH) models have successfully been implemented and just like stochastic volatility models, further extended by considering jumps in the underlying, non-Gaussian innovations or both (e.g. Christoffersen et al., 2006; Duan et al., 2006; Christoffersen et al., 2012). Of particular interest to us is the model of Heston and Nandi (2000). It is the only GARCH model which provides a quasi closed-form solution for the valuation of European options what makes it very convenient to apply in an empirical analysis.

Unlike in continuous-time models, however, with only a single source of uncertainty, volatility in GARCH models is time varying but not stochastic. Moreover, historical information is incorporated in a GARCH framework, either through a return based estimation or jointly with option data. While many of these works lead to a reduction in pricing errors compared to the constant volatility model of Black and Scholes (1973), mispricings are still an issue and several other questions remain unanswered. According to (e.g. Bates, 1996), for instance, a key aspect in option pricing is the lack of consistence in physical and risk-neutral distributions and an inadequate understanding of their relation. Furthermore, due to the inherently different model structure, a direct comparison between stochastic volatility and GARCH models is difficult and rarely conducted. Yet, this does not mean that understanding the different option pricing models and comparing their strengths and weaknesses should be neglected.

Accurately determining the value of a contingent claim is essential and highly relevant for a number of reasons. Through arbitrage strategies profits can be made and correctly hedged positions help with managing risks and minimizing losses. Motivated by increasing numbers of derivatives and ever larger positions, as well as regulatory pressure in the aftermath of the financial crisis, the objective of this thesis is thus to thoroughly investigate these issues. For this purpose, we focus on an in-depth analysis of the Heston and Nandi (2000) model and ultimately evaluate its performance. Our comprehensive analysis also encompasses the underlying GARCH dynamics and the impact of return shocks on our model. To better assess the results we include two benchmark models, namely the Black-Scholes-Merton (BSM) model and the Heston (1993) model. These are intended to give indications of performance limits to the down- and upside and therefore a range of realistic outcomes. Furthermore, the most likely hard to beat Heston (1993) model offers a closed-form solution as well. To tackle the central issue named by Bates (1996) we include an extension to our investigated GRACH model proposed by Christoffersen et al. (2013). Since we rely on the mapping of parameters between the physical and risk-neutral measures, the relation between these two is particularly relevant.

In contrast to Heston and Nandi (2000), who use Standard & Poor's 500 (S&P500) data from the early nineties, we rely on a more recent and much longer sample of German data. Due to changed market characteristics our estimation should yield different parameters. We test these for significance what hence should allow for additional insight. The GARCH models are estimated sequentially, beginning with returns. Once all parameters are obtained and all remaining models are calibrated to multiple cross-sections of data, we conduct our out-of-sample analyses on two separate samples. To highlight the importance of the objective function, we do so for two different loss functions and investigate the consequences. While the higher parameterized models should fit the data better, we expect the differences to converge in the forecasting period.

The remainder of this thesis is structured as follows: In Chapter 2 we introduce the fundamental principles of option valuation, including stylized facts and risk-neutral valuation. Furthermore, we describe the investigated models and highlight relevant characteristics. Chapter 3 contains the empirical analysis, where we begin by estimating the parameters, which are then used to compare fit and forecasting performances of the models in-sample and out-of-sample. Chapter 4 concludes.

2 Volatility Dynamics and Model Implementation

Chapter 2 lays the necessary foundation for the empirical analysis. Aside from explaining the fundamentals of option theory, we will gain an understanding of relevant concepts in option valuation, such as the risk-neutral valuation theory and its relationship to actually observed market data. We will also explore some statistical properties of asset returns, which an option pricing model should be able to capture. Furthermore, it is necessary to distinguish between discrete-time and continuous-time trading, since there are some important differences. Once the theoretical framework is set, the models analyzed in this work are explained in detail, starting with the well-established model of Black and Scholes (1973) and Merton (1973) and the Heston (1993) model. These are intended to serve as benchmarks for the central model of this thesis, the Heston and Nandi (2000) model. An extension to which, along the lines of Christoffersen et al. (2013), introducing an independent variance risk premium, is investigated as well.

2.1 Theoretical Framework

Since we are attempting to determine the value of financial options, it is sensible to give a brief explanation about what an option actually is. Generally, an option is a derivative, which means its price is derived from the price of another asset or a quantity such as an interest rate. While there are countless of differently structured derivatives and combinations of these, for options, most can be broken down to the plain vanilla structure of a call or a put. As the name itself already states, an option, or contingent claim, gives its holder the right to a buy or sell a certain underlying asset to previously agreed terms. The writer of the option acts as the counterparty and is obliged to deliver or accept the underlying at a certain price, the strike price. Depending on the nature of the contract, it can either be exercised at any time until expiration (American-style), or only be executed at a prespecified date (European-style). For option valuation the second case tends to be the simpler one, especially when dividends come into play. Contingent on not only the underlying S itself and its volatility, but also on the agreed strike price K, the risk-free rate of return and the time until expiration, the terminal payoff of a European call option at time $t = T$ is given by $C_T = [S_T - K]^+$. As this quantity is always larger or equal to zero, a call option can be seen as an insurance for an investor looking to buy an asset. At most she will lose the initial premium paid for the contract, should the asset fall in price. Similarly, a put option, $P_T = [K - S_T]^+$, offers protection to an investor looking to sell the asset underlying the option, since any potential loss is compensated for by the derivatives contract. While hedging strategies are a sensible way to apply options, one can also use them to speculate. In any scenario, however, it is important to determine the correct value of the option, which is composed of the intrinsic value and the time value.

For the purpose of option valuation, we will take a closer look at a number of different types of models. These models are designed to explain the non-linear relationship between the market price of an option and movements of the underlying, as well as certain variables. What all analyzed models have in common is their affine structure. This means a closed-form, or quasi closed-form, solution is available. While this is very convenient, it is not the usual case. Most option pricing models rely on rather time intensive Monte Carlo simulations to obtain results, which is why relatively few studies analyze non-affine models. Exceptions are Aït-Sahalia and Kimmel (2007), who analyze a more richly parameterized model, nesting a discrete-time, as well as the Heston (1993) model and Christoffersen et al. (2010b). Christoffersen et al. (2010b) on the other hand estimate different, but equally parsimonious, non-affine models and find substantial improvements over affine models, indicating that the convenience of analytical tractability comes at a price. Hsieh and Ritchken (2005) compare the non-affine model of Duan (1995) and the affine model of Heston and Nandi (2000). Despite doing very well, their empirical evidence also shows that the closed-form solution is outperformed, especially for shorter maturities and generally for out-of-the-money (OTM) contracts.

Irrespective of the chosen structure, any model should be able to capture various essential statistical properties of asset returns. These properties, common across different asset classes and time horizons, are labeled “stylized facts”. The asymmetry between gains and losses, i.e., large drawdowns, but not equally large upward movements in stock prices, is one example of a stylized fact. It occurs due to negatively skewed returns. While less pronounced for individual stocks, the effect is especially observed for daily returns on stock market indices (Munk, 2013, P. 14). The resulting distribution gives rise to frequent, but small, positive returns and occasional larger losses. Additionally, heavy tails and an excess kurtosis lead to non-normally distributed returns.

Another well-known attribute, termed the “leverage effect”, was originally observed by Black (1976). Related to the asymmetry of returns, it states that negative returns are mostly followed by larger increases in volatility compared to equally large positive returns. While Black (1976) related this effect to the indebtedness of a firm, reasoning that negative returns reduce equity but not debt and therefore risk of default, he also realized that this is not sufficiently explaining the actual asymmetry observed (see also Hentschel, 1995). Following works, most notably by Christie (1982), French et al. (1987), Nelson (1991) and especially Schwert (1989), who present additional evidence for a negative correlation between return volatility and equity value, also document these findings. The time varying volatility itself tends to be positively autocorrelated. Therefore, a period of high volatility is more likely to be succeeded by another period of high volatility, than it is not. This volatility clustering is another important statistical property. A list of further relevant stylized facts that should be considered when attempting to price contingent claims can be found in, for instance, Cont (2001).

What should not only be considered, but is an absolute necessity in option pricing, is the assumption of no-arbitrage. Given the possibility of a perfect hedge to future price fluctuations, it is possible to value contingent claims using the arbitrage pricing theory. The arbitrage pricing theory implies the existence of a risk-neutral probability measure, under which all assets have the same expected rate of return, equal to the risk-free rate. Furthermore, in an arbitrage free world, the individual risk preferences of investors do not affect valuation decisions (Delbaen and Schachermayer, 2006). Under such a risk-neutral probability measure, or equivalent martingale measure, the price of a European type option equals its discounted expected future payoff (Cox and Ross, 1976; Ross, 1976). For a call option it is given by

$$C_t = \exp\left(-r\left(T-t\right)\right) \mathbb{E}^{\mathbb{Q}}\left[S_T - K\right]^{+}. \tag{2.1}$$

Equally, the value of a European put option can be obtained by replacing the expected payoff accordingly, with $\mathbb{E}^{\mathbb{Q}}\left[K - S_T\right]^{+}$. Furthermore, in asset pricing the absence of arbitrage guarantees the existence of a non-negative pricing kernel, linking risk-neutral to actual probabilities and therefore reconciling physical with risk-neutral return distributions (Christoffersen et al., 2013).

While this works in a continuous-time environment, option prices do not follow from arbitrage-free pricing, as suggested for example in Harrison and Pliska (1981), in discrete time steps. The argument of a perfect hedge does not hold, since returns and volatility are not perfectly correlated in a non continuous-time model. Therefore, in order to obtain a risk-neutral valuation relationship in the discrete-time setting, it is necessary to build, for instance, on the approach of Rubinstein (1976) and Brennan (1979), who are able to derive preference-free option pricing equations in a discrete-time trading model (Câmara, 2003). They do so by applying general equilibrium arguments, which, unlike in the partial equilibrium approach, do not solely rely on the absence of arbitrage. Instead, investors' preferences play an important role (Musiela and Rutkowski, 2009, P. 14). The resulting risk-neutral valuation relationship is aligned with the assumption of risk-neutral preferences among investors and leads to the same expected rate of return for all securities. Just as in the continuous case, this is the risk-free rate.

For the purpose of a better interpretation the risk-free rate is set to zero in all following calculations. This is not necessarily a result-distorting assumption. For relatively short times to maturity, interest rates do not have a big impact on option prices. This should especially hold true in a virtually zero percent rate environment, with German government bonds of up to ten years temporarily even in slightly negative territory. Additionally, dividends are excluded as well. Since the underlying considered in this work, the German stock index DAX, is a total return index, its spot price already incorporates dividend payments. Any equation or formula given, will therefore not mention a dividend variable.

2.2 The Black-Scholes-Merton (1973) Model

Striving to obtain and benefit from a correct valuation formula, the first meaningful attempts in option pricing, dating back to at least 1900, are attributed to Louis Bachelier, who used equilibrium type arguments in his work. Yet, it wasn't until about seventy years later that the pioneering formula developed by Black and Scholes (1973) and further extended for constant continuous dividends by Merton (1973) sparked a vast amount of research in the field of option pricing. Their noble prize winning works also lead to a rapid expansion in the derivatives market. Given a set of arguably reasonable "ideal conditions", Black and Scholes (1973) derive a closed-form solution for the pricing of European style options. These conditions include the absence of arbitrage, perfectly divisible and readily available quantities of the underlying, zero taxes and no transactions costs. A further requirement is the possibility of continuous trading and the ability to borrow and lend arbitrarily at the theoretical, constant risk-free rate.

The stochastic process driving the price of the risky underlying in the BSM model is a geometric Brownian motion and is given by

$$dS_t = \upsilon S_t dt + \sigma S_t dW_t, \tag{2.2}$$

where W_t is a Wiener process, or standard Brownian motion. The annualized standard deviation of the instantaneous return $\frac{dS_t}{S_t}$, the instantaneous return volatility, is given by σ. υ describes the annualized drift of the lognormally distributed underlying S_t and is, just like the diffusion σ, assumed to be constant in the BSM world. This very restrictive assumption, that volatility is uncorrelated with spot returns, makes it impossible for the BSM model to capture the skewness effects observed on the markets.

Under the introduced assumptions, the value of the option now solely depends on the price of the underlying, the time until expiration of the contract and a set of constant variables (see Appendix A). The well-known option pricing formula can be obtained by employing delta hedging arguments and solving the resulting partial differential equation (A.8), using the terminal boundary condition $C_T = [S_T - K]^+$. This yields

$$C_t = S_t \Phi\left(d1\right) - K \exp\left(-r(T-t)\right) \Phi\left(d2\right), \tag{2.3}$$

where $\Phi\left(\cdot\right)$ denotes the cumulative standard normal distribution function

$$\Phi\left(a\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a} \exp\left(-\frac{1}{2}y^2\right) dy, \tag{2.4}$$

with a replaced by $d1$ and $d2$ respectively. These are given by

$$d1 = \frac{\log\left(\frac{S_t}{K}\right) + \left(r + \frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}, \tag{2.5}$$

$$d2 = d1 - \sigma\sqrt{T-t}. \tag{2.6}$$

Plugging in all values gives the price of a European call option at time t, where t is usually set to zero to get the price of the contract for today, expiring after T years in time. Similarly, the value of a put option can be calculated through

$$P_t = K \exp\left(-r(T-t)\right) \Phi\left(-d2\right) - S_t \Phi\left(-d1\right). \tag{2.7}$$

In both cases the spot and strike price are multiplied by some probability, depending on $d1$ and $d2$. Of these, $\Phi(d2)$ gives the risk-neutral probability that the option will expire in the money and therefore be exercised. Slightly less straightforward, $\Phi(d1)$ expresses the probability by which the present value of the underlying, contingent on finishing in the money, is larger than the current spot price (Nielsen, 1993). To price a European put option, one can also rely on the put-call parity

$$P_t = C_t - S_t + K \exp\left(-r\left(T-t\right)\right). \tag{2.8}$$

Given the price of a call option on the same underlying, with equal strike and expiry date, the price of the put can be inferred.

As can be seen in equations (2.3) and (A.8), the option price in this model is not influenced by any remaining parameter for the market price of risk or the attitude of investors towards risk. This is due to the key idea of the BSM model, that there is only one source of risk affecting the price of the underlying (Munk, 2013, P. 518). Since one can replicate the option through a portfolio of the underlying and a risk-free asset, or alternatively, create a risk-free portfolio through a combination of the underlying and the option, the return has to be equal to the risk-free rate. While this only holds true for that very moment in time, under the initially stated assumptions, dynamic adjustments are easily applicable.

Nonetheless, as groundbreaking the results of Black and Scholes and Merton were, they are and were not very realistic, since the model relies on several restrictive assumptions. Returns are not necessarily following a normal distribution and therefore the BSM model fails to capture a number of important stylized facts. Additionally, implied volatility, which is the volatility one has to use as an input in equation (2.2) to obtain the actual market price of an option, is found to vary greatly with strike price and time left until maturity. If the BSM model were a correct representation of market dynamics, implied volatility would be the same for all options written on an underlying.

2.3 The Heston (1993) Model

Heston (1993) relaxes the assumption of constant volatility by introducing a second Brownian motion, intended to describe the time-varying fluctuations. Despite some remaining well-known shortcomings, the application and usage of the Heston (1993) stochastic volatility model, henceforth “Heston model”, is ubiquitous. This is usually attributed to its relative simplicity and comparably low computational cost due to the quasi closed-form solution. It can also be related to a lack of superior competing attempts that would consistently outperform the Heston model in all dimensions, especially out-of-sample. One major shortcoming usually cited is that even though arbitrary, constant skewness and kurtosis are not sufficiently explaining deep out-of-the-money option prices with only days left to maturity (e.g. Nandi (1996); Moyaert and Petitjean (2011)). Despite this criticism, the Heston model is a huge improvement over the BSM model. It not only allows for fat tails by raising kurtosis and therefore increasing the price of options far from the strike, but also for skewness, to control better for differences between options written with lower and higher strikes. Among the most notable contributions to tackle the issue of short-term far out-of-the-money options are proposals to add Poisson processes to model jumps. Either in return dynamics of the underlying (Bates, 2000), or alternatively jumps in volatility, as for instance by Eraker (2004). Another suggestion includes a second variance process to increase flexibility with regard to the term structure, which allows the independent modelling of slope and level of the volatility (Christoffersen et al., 2009).

In the classical Heston (1993) framework though, the price of the underlying is assumed to follow the diffusion in equation (2.9). Given that volatility follows an Ornstein-Uhlenbeck process one can apply Ito's Lemma to obtain the variance process, which can then be rewritten as the square root diffusion process in equation (2.10). The stochastic differential equations are

$$dS_t = \upsilon S_t dt + \sqrt{\nu_t} S_t dW_t^{(1)}, \tag{2.9}$$

$$d\nu_t = \kappa\left(\theta - \nu_t\right) dt + \varrho\sqrt{\nu_t} dW_t^{(2)}, \tag{2.10}$$

with

$$dW_t^{(1)} dW_t^{(2)} = \rho dt, \tag{2.11}$$

where S_t is the price of the underlying at time t with drift υ and ν_t is its corresponding instantaneous variance. The two sources of randomness driving the underlying and the variance are the Wiener processes $W_t^{(1)}$ and $W_t^{(2)}$. The time-invariant parameters κ, θ, ϱ and ρ are commonly referred to in literature as the rate of mean-reversion, the long-term mean of variance and the volatility of volatility, with ρ representing the constant correlation between the two independent Wiener processes. In equity markets, the usually observed case of $\rho < 0$ implies that a negative return tends to lead to an increase in volatility and the probability density is skewed to the left. Vice versa, for $\rho > 0$ volatility would drop along with a negative underlying return, therefore exhibiting a fatter tail to the right. It is important to note that there is also a state variable, the initial variance ν_0, which can either be estimated along with the other parameters or, for instance, be filtered from returns of the underlying. The process stated in (2.9) is almost the same price process that drives the infamous Black-Scholes model, with the exception of a time varying volatility. Setting the parameter affecting the kurtosis of the spot returns ϱ in (2.10) equal to zero, the Heston dynamics approach those of Black and Scholes (1973).

In order to price contingent claims, in this case plain vanilla options, it is required to derive the risk-neutral dynamics of (2.9) and (2.10), relating the value of the option to the value of its underlying and other variables (Christoffersen et al., 2010a).

For $x_t = \log(S_t)$ the risk-neutral dynamics in the Heston model are given by

$$dx_t = \left(r - \frac{1}{2}\nu_t\right) dt + \sqrt{\nu_t} dW_t^{*(1)}, \tag{2.12}$$

$$d\nu_t = \kappa^* \left(\theta^* - \nu_t\right) dt + \varrho \sqrt{\nu_t} dW_t^{*(2)}, \tag{2.13}$$

where $\kappa^* = \kappa + \lambda$ and $\theta^* = \frac{\kappa\theta}{\kappa+\lambda}$ (see Appendix B). $\lambda(S, \nu, t)$ is the market price of volatility risk, subject to the price and volatility of the underlying, as well as the time t. Since the market price of volatility risk is not an observable quantity and specifying it, as well as estimating it accurately, proves to be a difficult task, it is convenient that the risk-neutral formulation does not require knowledge of λ. On the other hand, it is not possible to identify λ in this model, since as long as only option data are used, the volatility risk premium cannot be isolated.

Initially proposed in a stochastic volatility model by Stein and Stein (1991), Heston (1993) takes the inverse Fourier transform to obtain an analytical valuation formula for European options. Following the characteristic function (CF) approach the value of a European type call at time t is given by

$$C_t = S_t \Pi_1 - K \exp(-r\tau) \Pi_2, \tag{2.14}$$

with

$$\Pi_1 = \frac{1}{2} + \frac{1}{\pi} \int_0^\infty \Re e \left(\frac{\exp(-\imath\phi \log(K)) f^*(x, \nu, \tau, \phi - \imath)}{\imath\phi S_t \exp(r\tau)} \right) d\phi, \tag{2.15}$$

$$\Pi_2 = \frac{1}{2} + \frac{1}{\pi} \int_0^\infty \Re e \left(\frac{\exp(-\imath\phi \log(K)) f^*(x, \nu, \tau, \phi)}{\imath\phi} \right) d\phi, \tag{2.16}$$

where $\Re e(\cdot)$ indicates the Real part of the expression in brackets, $\imath$ is the imaginary unit and $\tau = T - t$. The CF in risk-neutral terms is

$$f^* (x, \nu, \tau, \phi) = \exp\left(A(\tau, \phi) + B(\tau, \phi)\, \nu_0 + \imath\phi x\right), \tag{2.17}$$

$$A(\tau, \phi) = r\phi\imath\tau + \frac{\kappa^*\theta^*}{\varrho^2}\left(\left(\kappa^* - \varrho\imath\rho\phi - d\right)\tau - 2\log\left(\frac{1 - g\exp(-d\tau)}{1-g}\right)\right), \tag{2.18}$$

$$B(\tau, \phi) = \frac{\kappa^* - \varrho\imath\rho\phi - d}{\varrho^2}\left(\frac{1 - \exp(-d\tau)}{1 - g\exp(-d\tau)}\right), \tag{2.19}$$

$$g = \frac{\kappa^* - \varrho\imath\rho\phi - d}{\kappa^* - \varrho\imath\rho\phi + d}, \tag{2.20}$$

$$d = \sqrt{\left(\kappa^* - \varrho\imath\rho\phi\right)^2 + \varrho^2\left(\phi^2 + \phi\imath\right)}. \tag{2.21}$$

Inverting the CF of the logarithm of the spot price (2.17) according to equations (2.15) and (2.16) yields the desired probabilities. These integrands have to be evaluated numerically. However, this does not pose much of a problem, since both functions decay rapidly and smoothly converge to zero. Therefore, using appropriate integration methods allows for a fast calculation.

Depending on the chosen formulation of the CF, the time left until maturity and the selected parameter set, difficulties can occur when calculating option prices. Specifically, the evaluation of the complex logarithm in equation (2.18) is troublesome when the Feller Condition is violated, i.e., $(2\kappa^*\theta^* < \varrho^2)$.[1] While always non-negative, for certain combinations of κ^*, θ^* and ϱ, the Cox et al. (1985) (CIR) process in equation (2.13) can reach zero with a positive probability, therefore leading to discontinuities and rendering any result meaningless, since option prices are calculated erroneously.[2] In 1999, Schöbel and Zhu were the first to document this problem, which also occurs in the original CF of Heston (1993) and a plethora of other work, relying on Heston's formulation. Aware of the phenomenon, Kahl and Jäckel (2005) tackle the issue and propose a solution which keeps track of the number of branch cuts, i.e., crosses of the negative real axis by the complex power function in the terms of equation (2.17).

Another expression, now conventionally accepted, evolved around the works of Schoutens et al. (2003), Albrecher et al. (2006), Gatheral (2006) and Lord and Kahl (2006). Lord and Kahl (2006) acknowledge the preceding conjecture of Gatheral (2006) and prove, that given certain parameter restrictions, this formulation does not suffer from the branch cut problem in the complex plane (see Appendix C). Simultaneously and independently, Albrecher et al. (2006) provide a universal proof, showing that this also holds true for the full dimensional and unrestricted parameter space.[3] Furthermore, seemingly unaware,

[1]Since the complex logarithm in general is not continuous along the negative real axis.

[2]For $2\kappa^*\theta^* \geqslant \varrho^2$ Cox et al. (1985) note that the upward drift is large enough to keep the variance from reaching zero at all times.

[3]See also Lord and Kahl (2010) for a summary and adaptions to other models.

since they do not provide any explanation for their choices, other authors use the correct terms or alternative formulations in earlier works, leading to continuous expressions. Most notably Bakshi et al. (1997) and Duffie et al. (2000).

Yet, even more than 20 years after its publication and countless of contributions later, the properties of the Heston model are not entirely understood. While the quasi-analytical closed-form solution has already been a part of Heston's seminal work and issues regarding its continuity have gradually been fixed, there is still an extensive ongoing research. For instance, Andersen (2007) introduces and analyzes discretization schemes for Monte Carlo simulation, which is highly relevant for path-dependent securities. In another, very recent, contribution Cui et al. (2015) derive the analytical gradient of the option pricing function, allowing a significantly faster ($10x - 15x$) calibration of the Heston model, compared to a numerical gradient. This could especially prove useful when applied to larger datasets or in high-frequency trading.

Benchmarking a variety of CFs for the Heston model and attempting to optimize their performance with regards to a fast and precise calculation, ultimately led to the formulation of the CF in (2.17) in our case. Among others compared were, aside from the original and the chosen formulation, alternative approaches such as the ones proposed in Duffie et al. (2000) or Attari (2004), who use the imaginary part or a combined integrand, respectively.

2.4 The Heston & Nandi (2000) Model

Unlike the two previously described models, who have a number of similarities in their structure, the Heston and Nandi (2000) model does not rely on the assumption of continuous-time trading. All actions in a discrete-time model take place in a finite number of steps. Since official markets do not offer the opportunity of uninterrupted trading outside the regular hours, this seems to be a more plausible representation of reality. While one could argue that over-the-counter markets mitigate this problem, there is still the situation of non-zero transaction costs. And despite small spreads that are present in larger and more liquid markets, including the German blue chip index DAX, even large institutions face noticeable transaction costs (Lesmond et al., 1999). Liu and Loewenstein (2002) for instance note that even with small transaction costs optimal behavior changes dramatically. They find that depending on the market situation, transaction costs can lead to suboptimal buy-and-hold strategies. This is in line with Gârleanu and Pedersen (2013), who also report that transaction costs lead to a deviation from the otherwise considered optimal result and should not be neglected.

2.4.1 GARCH Processes in Option Valuation

The Heston and Nandi (2000) (HN) model rests on the main assumption that the log-spot price follows a particular GARCH process, given in equations (2.25) and (2.26). However, before introducing the HN model in detail, we first take a closer look at the theory underlying its alternative type of modelling approach. Since asset returns tend to exhibit little to no autocorrelation, previous concepts mainly relied on random walks to model prices. The absence of autocorrelation in returns alone, however, does not imply independence (Cont, 2001). As Lo and MacKinlay (2008, P. 17-46) note, asset returns are typically not independent and therefore not in line with the popular random walk hypothesis. Initially proposed by Engle (1982) as Autoregressive Conditional Heteroskedasticity (ARCH) model, with the intention to describe inflationary uncertainty in the United Kingdom, a new class of processes would prove helpful in tackling this issue. Rejecting the assumption of constant variance, Engle's approach allows the conditional variance to fluctuate over time, while leaving the unconditional variance constant. The following generalization suggested by Bollerslev (1986), adding lagged conditional variances, has since been implemented in numerous works in one way or another. In its general form the original GARCH(p, q) model is given by

$$h_t = \omega + \sum_{j=1}^{p} \beta_j h_{t-j} + \sum_{j=1}^{q} \zeta_j \varepsilon_{t-j}^2, \tag{2.22}$$

with p representing the number of autoregressive terms, i.e., lags of the conditional variance h_t and q the number of moving average terms considered. The residual ε_t has the same functional form as in the ARCH model and can be decomposed as

$$\varepsilon_t = z_t \sqrt{h_t}, \tag{2.23}$$

where $z_t \overset{iid}{\sim} \mathcal{N}(0, 1)$ and therefore $\varepsilon_t \sim \mathcal{N}(0, h_t)$, conditional on past information. Compared to the ARCH process the GARCH, even in the simplest $(1, 1)$ formulation, considers past conditional variances as well and does not suffer from an abrupt cutoff after q lag terms. Instead it includes all past lags, like an infinite order ARCH, with geometrically decreasing influence, back to the initial period (Engle and Ng, 1993). This structure, with the particular parameterization of $p = q = 1$, has independently been proposed by Taylor in 1986. In the case of $p = 0$, the GARCH process reduces to an ARCH(1) process, which is an exception, since this process is markovian. For all other parameter combinations, the GARCH model is non-markovian, as it incorporates past information. This differs from the usual approach in option pricing, which assumes the underlying to follow a memoryless diffusion process (Duan, 1995). Limiting both, p and q to zero, all that's left is a simple white noise process, equal to the homoskedastic process in the BSM model.

Building on the revolutionary works of Engle (1982) and Bollerslev (1986), a large variety of extensions has been proposed since. The original formulation, which only contains squared innovations, does not distinguish between positive and negative shocks. Therefore, even though accounting for the magnitude of a move, the conditional variance shows the same response, irrespective of the sign of the shock. Yet, observations where negative returns increase variance by more than similarly sized positive ones do, are common. Another desirable property a model should be able to capture is the convergence to a long-run unconditional variance. While the GARCH model of Bollerslev (1986) allows for stochastic, mean-reverting volatility dynamics, it does not account for the asymmetry, required to model the leverage effect (Barone-Adesi et al., 2008). Several solutions, such as the exponential GARCH (EGARCH) by Nelson (1991), the non-linear asymmetric GARCH (NGARCH) of Engle and Ng (1993), or the GJR-GARCH, named after Glosten et al. (1993), do especially well in creating the necessary asymmetry. Heston and Nandi (2000) propose a new formulation for the conditional variance, very similar to Nelson's NGARCH. We will call this the Heston & Nandi GARCH (HNGARCH). Unlike in the GJR or the EGARCH formulation, the asymmetry does not stem from different slopes on either side of the axis of ordinates, but a shift of the so called News Impact Curve (NIC).

The mentioned GJR-GARCH$(1,1)$ model is given by

$$h_t = \omega + \beta h_{t-1} + \zeta \varepsilon_{t-1}^2 + \vartheta S_{t-1}^- \varepsilon_{t-1}^2, \tag{2.24}$$

where $S_{t-1}^- = 1$ for $\varepsilon_{t-1} < 0$ and $S_{t-1}^- = 0$ otherwise.

Computing the News Impact Curves for three different approaches, by calculating the respective equations for different error terms, we will gain some insight regarding the effect of a past return shock on the conditional variance. Given volatility is at its long-run level, the NIC measures exactly that. Yet, one has to be careful when interpreting the graphs, since average pricing parameters are used to establish an understanding of the dynamics. These, with the exception of the auxiliary parameters ζ and ϑ, are obtained through the estimation undertaken in Chapter 3.2. ζ is set to $\alpha \times 5694.40 = 0.0465$[4] and ϑ is adjusted to $(\gamma + \mu) \times 10^{-7} = 1.2405 \times 10^{-5}$, in order for the models to match each other's dimensions. The remaining parameters, which will be estimated and explained later, are $\alpha = 8.1688 \times 10^{-6}$, $\beta = 0.8063$, $\gamma = 121.56$, $\omega = 3.7568 \times 10^{-6}$, $\mu = 2.49$ and $\sqrt{h_t \times 252} = 21.04\%$. Note, to simplify interpretation, the vertical axis in Figure 1 shows the annualized conditional volatility rather than the daily conditional variance, as in Engle and Ng (1993).

It can immediately be spotted, that the HNGARCH is not centered at the origin. Instead, for positive γ, its minimum lies to the right, with the extent of the shift depending on the standard deviation. For an innovation of $z_{t-1} = \gamma \sqrt{h_{t-1}}$ the squared term in

[4]Since multiplying α by 5694.40 matches the ARCH terms of the models when γ is 0.

equation (2.26) is zero, therefore minimizing the curve. This way of analyzing the impact of shocks was first applied by Pagan and Schwert (1990) and later popularized by Engle and Ng (1993), who came up with the term News Impact Curve. Visualizing the effect of innovations also shows that the HNGARCH(1, 1) is simply the GARCH(1, 1) shifted to the right. Both are symmetric functions, despite the HNGARCH model itself, due to the shift, being asymmetric. The differing GJR approach on the other hand is centered at the origin but exhibits a much steeper slope for negative innovations, accounting for asymmetry and therefore the leverage effect in this fashion. Using the same parameters, it is equal to the GARCH(1, 1) model for positive innovations.

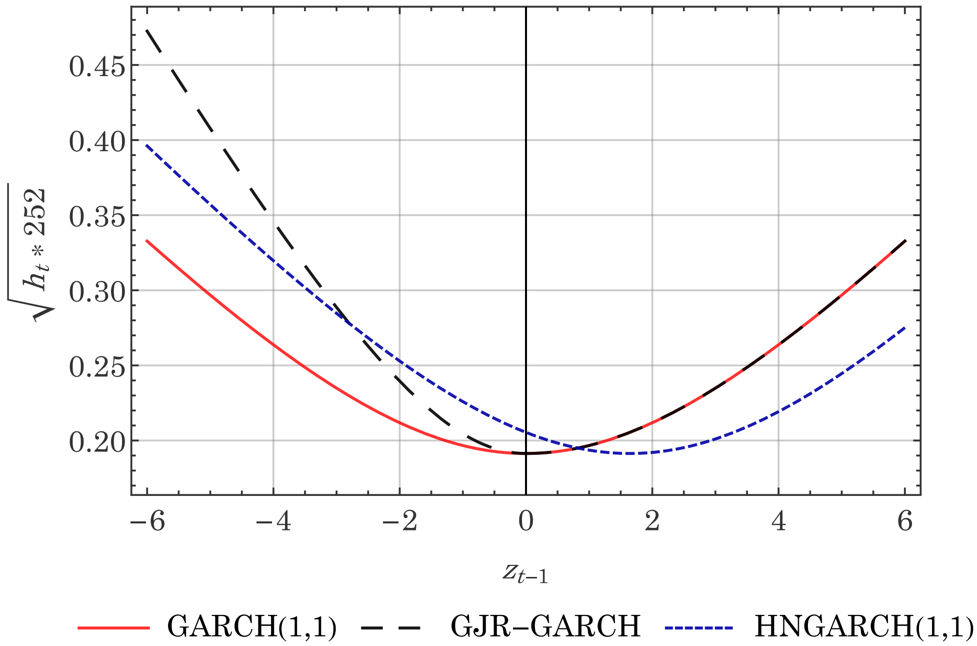

Figure 1 – Stylized news impact curves, indicating the response to differently sized return innovations, for the GARCH(1, 1) model, the GJR-GARCH(1, 1) model and the single lag case of the model of Heston and Nandi (2000).

Engle and Ng (1993) report that the GJR-GARCH model captures volatility dynamics very well, even in extreme scenarios. For reasonably scaled shocks they note that there is not much of a difference between the GJR and the simple GARCH(1, 1) formulation. Of those mentioned here, the HNGARCH offers the only solution which allows the conditional variance to get lowest when slightly positive return innovations occur. While this specification tends to work well for smaller shocks, Barone-Adesi et al. (2008) find that the GJR approach, especially with a non-normal but instead historically filtered distribution of innovations, performs much better when applied to market data. They also show that the normality assumed in, for instance, the Heston and Nandi (2000) return innovations is not sufficient to explain observed shocks. These realized shocks, especially large negative ones, imply the arrival of significant news and often exceed a value of -3, what is rarely the case for a standard normal distribution ($\approx 0.14\%$). As can be seen from Figure 1, the asymmetry introduced by the shift becomes more and more negligible, leading to similar responses for large positive or negative innovations. Increasing the parameter γ to account for this effect simply shifts the HNGARCH$(1,1)$ further to the right, without changing the slope. As can be deduced from the last term in equation (2.24), a similar change of ϑ in the GJR model steepens the curve, yet, due to the indicator function, this only affects negative innovations.

2.4.2 Model Properties and Risk-Neutral Dynamics

Just like for continuous-time models, barely any closed-form solution for GARCH models exists, which is why most approaches rely on simulation. An exception, building on the works of Engle and Mustafa (1992), Amin and Ng (1993) and Duan (1995), is the model of Heston and Nandi (2000). Heston and Nandi (2000) develop an almost closed-form valuation formula for European type options. In their affine discrete-time model the physical return and variance dynamics for the first-order case are

$$\log(S_t) = \log(S_{t-\nabla}) + r + \left(\mu - \frac{1}{2}\right) h_t + \sqrt{h_t} z_t, \tag{2.25}$$

$$h_t = \omega + \beta h_{t-\nabla} + \alpha \left(z_{t-\nabla} - \gamma \sqrt{h_{t-\nabla}}\right)^2. \tag{2.26}$$

For our situation of a fixed daily interval ∇ is equal to 1. Therefore, r represents the daily continuously compounded interest rate. z_t is the only source of randomness in this model and follows a standard normal distribution, while α, β, γ, μ and ω are the model parameters. μ is a linear equity risk premium, such that $\mu - \frac{1}{2}$ indicates the return premium per unit of risk, h_t. This functional form makes sure that for zero variance and therefore zero risk, equation (2.25) collapses to $R_t = \frac{\log(S_t)}{\log(S_{t-1})} = r$, i.e., the return of a period is equal to the risk-free rate over that time step, which prevents arbitrage. Furthermore, the model does not contain a separate parameter for a variance risk premium.

Since α determines the kurtosis of the distribution of log-returns, any positive α leads to a fat tailed and leptokurtic distribution. As we noted in Section 2.4.1, the parameter responsible for introducing asymmetry is γ, with a positive value leading to negative skewness. Setting the parameters α and β to zero and keeping ω strictly positive, yields again the BSM model, this time in discrete intervals. For the common case of $\alpha > 0$, however, ω is allowed to be zero in the HN model. As long as the variance persistence, which is given by

$$\beta + \alpha\gamma^2, \tag{2.27}$$

is smaller than 1, the unconditional variance will still be positive (Christoffersen et al., 2014). Should the sum in equation (2.27) be exactly equal to 1, this implies that the variance process in the GARCH$(1,1)$ case is integrated, which is also referred to as integrated

GARCH, or IGARCH, process (Duan, 1995). This special parameterization leads to non mean-reverting variance processes in the sense that past shocks remain relevant for any future forecast, which typically have infinite unconditional variances (Bollerslev and Engle, 1993). According to Bollerslev (1986), the GARCH$(1,1)$ model is only stationary for a persistence smaller than 1. His definition was further investigated and refined by Nelson (1990), who found that even the IGRACH process may still be stationary, depending on the parameter ω. As long as $\omega = 0$ the conditional variance converges to either zero, when $\beta + \alpha\gamma^2 < 1$, or infinity, when $\beta + \alpha\gamma^2 > 1$, almost surely.[5] For the case of the IGARCH, Nelson (1990) notes, that the conditional variance is simply a driftless random walk. Yet, for $\omega > 0$ the model is strictly stationary and the parameter ω can be interpreted as a positive drift. Assuming and limiting $\beta + \alpha\gamma^2$ to be smaller than 1 ensures covariance stationarity and therefore a finite mean and variance in our first-order case. It further allows to directly compute the conditional variance for time $t+1$, as a function of historical spot prices and the GARCH parameters known at time t. Solving equation (2.25) for z_t, plugging the result into equation (2.26) and simplifying yields

$$h_{t+1} = \omega + \beta h_t + \alpha \left(\frac{\log\left(S_t\right) - \log\left(S_{t-1}\right) - r - \left(\mu - \frac{1}{2} + \gamma\right) h_t}{\sqrt{h_t}} \right)^2, \tag{2.28}$$

what offers an important advantage over continuous-time stochastic volatility models. As we've learned in the previous sections, these require the filtering or re-estimation of the non-observable instantaneous volatility. Even though this does not pose a problem for just one cross-section of option prices, as done in Bakshi et al. (1997), the computational burden increases heavily for multiple cross-sections.

The unconditional, or long-run variance in the HN model is given by

$$\mathbb{E}\left[h_t\right] = \frac{\omega + \alpha}{1 - \beta - \alpha\gamma^2}, \tag{2.29}$$

where the denominator, $1-\beta-\alpha\gamma^2$, is simply 1 minus the persistence and can be interpreted as the variance mean-reversion. The variance of the conditional variance, as well as the covariance are

$$Var\left(h_{t+1}\right) = 2\alpha^2\left(1 + 2\gamma^2 h_t\right) \tag{2.30}$$

and

$$Cov\left(R_t, h_{t+1}\right) = -2\alpha\gamma h_t, \tag{2.31}$$

[5] With a probability of 1.

with the full derivation in Appendix D. As can be seen, the conditional variance is linear in past variance. For $\alpha > 0$ and $\gamma > 0$ the covariance is negative, indicating the desired effect of a negative correlation between spot returns and variance.

Reducing ∇ further, Heston and Nandi (2000) show that the process given in equations (2.25) and (2.26) is a discrete-time analog of the CIR square root process and that the diffusion process of Heston (1993) can be obtained as a limit. While this is an interesting property the practical relevance is known to be rather modest. As Corradi (2000) finds, a discrete-time model can have more than one continuous-time limit and vice versa. Depending on the methodology used, one can even obtain a continuous-time limit characterized by two independent, or at least non-perfectly correlated, Brownian motions, while the originating GARCH process only has a single source of innovation to begin with. Nelson (1991) and Duan (1997) for instance derive diffusion models with two sources of noise as the limit of a GARCH model. Kallsen and Taqqu (1998) on the other hand, using a different technique, yield a continuous-time model with only one innovation. By allowing the instantaneous variance to change only at the end of a given period, e.g. days, their approach makes it possible to price options based on the absence of arbitrage.

While this is not an option in the HN model, Heston and Nandi (2000) derive the risk-neutral dynamics by assuming that the value of a call option one period from expiration follows the Black-Scholes-Rubinstein formula. This central assumption is equivalent to the result of Duan (1995), who developed a GARCH option pricing model utilizing a generalized version of risk-neutralization, which he refers to as the locally risk-neutral valuation relationship. The risk-neutral dynamics for the HNGARCH$(1,1)$ model in equations (2.25)–(2.26) are

$$\log\left(S_t\right) = \log\left(S_{t-1}\right) + r - \frac{1}{2}h_t^* + \sqrt{h_t^*}z_t^*, \tag{2.32}$$

$$h_t^* = \omega + \beta h_{t-1}^* + \alpha\left(z_{t-1}^* - \gamma^*\sqrt{h_{t-1}^*}\right)^2, \tag{2.33}$$

with risk-neutral skewness parameter $\gamma^* = \gamma + \mu$, normal innovations $z_t^* \overset{iid}{\sim} \mathcal{N}(0,1)$ and an expected unconditional variance of

$$\mathbb{E}\left[h_t^*\right] = \frac{\omega + \alpha}{1 - \beta - \alpha\left(\gamma^*\right)^2}. \tag{2.34}$$

The value of a call option with strike K and expiry at $t = T$ can again be obtained through equation (2.14). Rearranging the formulae given in Heston and Nandi (2000) yields

$$\Pi_1 = \frac{1}{2} + \frac{\exp\left(-r\left(T-t\right)\right)}{\pi}\int_0^\infty \Re e\left(\frac{K^{-\imath\phi}f^*\left(x,t,T,\imath\phi+1\right)}{\imath\phi S_t}\right)d\phi, \tag{2.35}$$

$$\Pi_2 = \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \Re e\left(\frac{K^{-\imath\phi}f^*\left(x,t,T,\imath\phi\right)}{\imath\phi}\right)d\phi, \tag{2.36}$$

where Π_1 is the delta of the call value and Π_2 the risk-neutral probability of the price of the asset being greater than its strike at maturity. Both of these converge smoothly and quickly towards zero, what allows for an efficient numerical evaluation. The characteristic function can be calculated recursively and takes the log-linear form

$$f^* (x,\, t,\, T,\, \phi) = S_t^{\phi} \exp\left(A (t, T, \phi) + B (t, T, \phi)\, h_{t+1}^*\right), \tag{2.37}$$

$$\begin{aligned} A (t, T, \phi) = {} & A (t+1, T, \phi) + \phi r + B (t+1, T, \phi)\, \omega \\ & - \frac{1}{2} \log \left(1 - 2\alpha B (t+1, T, \phi)\right), \end{aligned} \tag{2.38}$$

$$\begin{aligned} B (t, T, \phi) = {} & \phi \left(\gamma^* - \frac{1}{2}\right) - \frac{1}{2} \left(\gamma^*\right)^2 + \beta B (t+1, T, \phi) \\ & + \frac{\frac{1}{2} \left(\phi - \gamma^*\right)^2}{1 - 2\alpha B (t+1, T, \phi)}, \end{aligned} \tag{2.39}$$

with terminal conditions $A(T, T, \phi) = 0$ and $B(T, T, \phi) = 0$ in the single lag case.

For this setup, the only parameter that differs in the option pricing dynamics is γ^*, which is due to the asset risk premium. All remaining parameters are the same for the physical and the risk-neutral distributions. While it is convenient to have such a unified framework, empirical studies find that the wedge created by just the risk premium, is not sufficient and leads to comparably poor performance. Pastorello et al. (2000), for instance, find persuasive evidence in their estimation of the Hull and White (1987) model for using option instead of return data. Chernov and Ghysels (2000) conduct a joint estimation of the Heston (1993) model, which shows similar results in favor of incorporating information from options. Similarly, Christoffersen and Jacobs (2004b) compare a number of differently parameterized GARCH models and point out the importance of the pricing kernel when relying solely on return data. While they find mixed results regarding improvements due to the inclusion of options in their estimation, they report very different parameters under the physical and risk-neutral measures. Heston and Nandi (2000), who include option as well as time-series data in their approach, equally obtain dramatically different parameter estimates between historical and joint estimation. One possible reason is pointed out by Barone-Adesi et al. (2008). They argue that the changing volatility leads to incomplete markets, what makes the replication argument of Black and Scholes (1973) invalid. As a result, investors would no longer price an option simply with a different drift, but would also take into account the changes in volatility.

2.5 The Christoffersen, Heston and Jacobs (2013) Model

A joint estimation, taking advantage of all the information contained in historical returns and in option data, is not an easy task. Especially identifying the latent volatility variable is troubling when using a stochastic volatility model. Thus, Christoffersen, Heston, and Jacobs (CHJ) (2013) develop a discrete-time model, which attempts to account for these matters. They argue that the HN model performs very well when fitting return or option data, but since different parameters are needed for both cases a joint estimation is problematic. By increasing and refining the wedge, i.e., the pricing kernel or relationship linking the physical and risk-neutral measure, CHJ (2013) attempt to improve the previously introduced HN model. Unlike Duan (1995) or Heston and Nandi (2000), who apply the power pricing kernel of Rubinstein (1976), CHJ (2013) propose a discrete-time analog of the pricing kernel implicit in the Heston (1993) model. For a constant variance it is as well equal to the monotonic power utility case of Rubinstein (1976) and Brennan (1979), but for stochastic volatility there are a number of implications. Their approach includes an additional parameter for the price of risk caused by fluctuations of the underlying, which needs to be estimated using option data. The total variance risk premium then consists of this new, independent parameter and another component, which is based on the covariance with equity risk. Furthermore, when the independent variance risk premium, ξ, is set equal to zero, the extended model nests the HN model as a special case. The pricing kernel, or stochastic discount factor, has the form

$$M_t = M_0 \left(\frac{S_t}{S_0}\right)^{\phi} \exp\left(\delta t + \eta \sum_{s=1}^{t} h_s + \xi \left(h_{t+1} - h_1\right)\right). \tag{2.40}$$

Leaving the physical process of equation (2.25) unchanged and applying the augmented pricing kernel leads to the following risk-neutral formulation

$$\log\left(S_t\right) = \log\left(S_{t-1}\right) + r - \frac{1}{2}h_t^* + \sqrt{h_t^*}z_t^*, \tag{2.41}$$

$$h_t^* = \omega^* + \beta^* h_{t-1}^* + \alpha^* \left(z_{t-1}^* - \gamma^* \sqrt{h_{t-1}^*}\right)^2, \tag{2.42}$$

with only the parameter $\beta^* = \beta$ remaining the same across both measures. z_t^* is still standard normally distributed, yet the new mapping is

$$h_t^* = \frac{h_t}{1 - 2\alpha\xi}, \tag{2.43}$$

$$\alpha^* = \frac{\alpha}{\left(1 - 2\alpha\xi\right)^2}, \tag{2.44}$$

$$\omega^* = \frac{\omega}{1 - 2\alpha\xi}, \tag{2.45}$$

$$\gamma^* = \left(\mu - 0.5 + \gamma\right)\left(1 - 2\alpha\xi\right) + 0.5. \tag{2.46}$$

Even though most of the parameters changed, this formulation still has the same structure as the Heston and Nandi (2000) model. Therefore, we can use the CF in equation (2.37) and the probabilities in (2.35) and (2.36) to conveniently obtain a price. It is also immediately apparent that for $\xi = 0$ the mapping of all parameters is identical to the HN model. The terms including ξ collapse, leading to $h_t^* = h_t$, $\alpha^* = \alpha$, $\omega^* = \omega$ and $\gamma^* = (\mu + \gamma)$.

In their empirical analysis CHJ (2013) find, that the log ratios of risk-neutral and physical densities exhibit a U-shape pattern. This pattern is remarkably stable over the 14-year period they observed and despite some noise, it is surprisingly close to a quadratic function. More importantly, their analysis shows that the pricing kernel is far from linear, which renders the monotonic kernels rather inadequate to explain the data. A suitable link should be able to capture the implicit U-shape, which is also supported by Bakshi et al. (2010). Solving equations (2.25) and (2.26) for z_t and $(h_t - h_{t-1})$ and plugging the results, together with $R_t = \log\left(\frac{S_t}{S_{t-1}}\right)$, into equation (2.40), it can be shown that after simplifying, the logarithm of the pricing kernel takes the following quadratic form

$$\log\left(\frac{M_t}{M_{t-1}}\right) = \frac{\xi\alpha}{h_t}\left(R_t - r\right)^2 - \mu\left(R_t - r\right) + \left(\eta + \xi\left(\beta - 1\right) + \xi\alpha\left(\mu - \frac{1}{2} + \gamma\right)^2\right) h_t + \delta + \xi\omega + \phi r, \tag{2.47}$$

with

$$\phi = -\left(\mu - \frac{1}{2} + \gamma\right)\left(1 - 2\alpha\xi\right) + \gamma - \frac{1}{2}, \tag{2.48}$$

$$\delta = -\left(\phi + 1\right) r - \xi\omega + \frac{1}{2}\log\left(1 - 2\alpha\xi\right), \tag{2.49}$$

$$\eta = -\left(\mu - \frac{1}{2}\right)\phi - \xi\alpha\gamma^2 + \left(1 - \beta\right)\xi - \frac{\left(\phi - 2\xi\alpha\gamma\right)^2}{2\left(1 - 2\alpha\xi\right)}, \tag{2.50}$$

where ϕ can be interpreted as the aversion to equity risk. For a negative independent variance premium, ξ takes positive values and the natural logarithm of the conditional pricing kernel exhibits the desired U-shape (see Figure 2). Interestingly, as Christoffersen et al. (2013) note, this is the case even though the pricing kernel is monotonic in each, returns and volatility. However, projecting the effect of the negative independent variance risk premium on the returns of the underlying yields a non-monotonic function. Since volatility not only increases for large negative returns but also for large positive ones, this is an important aspect of the model. The parameters used to generate Figure 2 are from the third column of Table 5 and the conditional variance is set equal to the variance of the entire return sample.

As can be seen, the log pricing kernel of the HN model is a linear, monotonic function. Its slope depends on the size of the asset premium but for a positive premium the line is always decreasing in returns and passing through the origin. Due to the positive log ratio

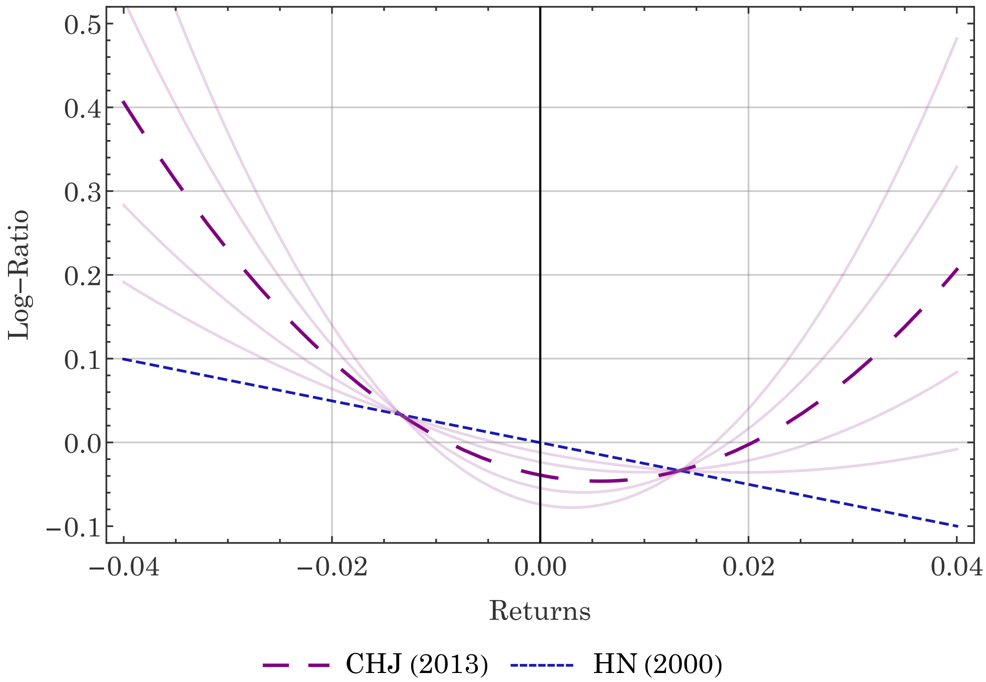

Figure 2 – Log ratio of the risk-neutral and physical density using the parameters from the third column in Table 5, for the freely estimated premium parameter, as well as the special case of zero independent variance risk premium ($\xi = 0$).

for negative returns the risk-neutral distribution has a fatter left tail. Gradually increasing the independent variance premium parameter produces a more and more pronounced U-shape. The dashed line, representing the in-sample estimates with $\xi = 4637$, indicates that the risk-neutral distribution has fatter tails than the physical distribution. For $\mu = 0$ the pricing kernel creates a perfectly symmetrical log ratio plot and setting all relevant risk premia equal to zero leads to identical empirical and risk-neutral probability measures.

Just like in the HN (2000) model the persistence in the CHJ (2013) model is given by

$$\beta^* + \alpha^* (\gamma^*)^2 . \tag{2.51}$$

However, due to the new mapping, specifically the scaling factor $\frac{1}{1-2\alpha\xi} = 1.0812$, the parameters α^*, γ^* and ω^* differ from their physical counterparts. Even with unchanged β^*, the effect on the persistence of shocks is not immediately apparent. While α^* and ω^* are always strictly larger when $0 < \xi < \frac{1}{2\alpha}$, the risk-neutral skewness parameter, γ^*, is negatively affected by an increase of the variance risk premium parameter. Generally, one has to be aware of the phenomenon that, unless restrictions are imposed, several values for ξ can lead to the same option price.

From equation (2.43) we know that the conditional variance implied by the risk-neutral process is larger in the model with an independent premium. The variance of the conditional variance and the covariance are, in unclear direction, affected as well with

$$Var\left(h_{t+1}^{*}\right)=2\left(\alpha^{*}\right)^{2}\left(1+2\left(\gamma^{*}\right)^{2}h_{t}^{*}\right) \tag{2.52}$$

and

$$Cov\left(R_{t},h_{t+1}^{*}\right)=-2\alpha^{*}\gamma^{*}h_{t}^{*}. \tag{2.53}$$

Relying on the parameter estimates from Table 5 again we can also analyze the NIC produced by the model with augmented pricing kernel. One weakness of the kernel implicit in the HN model was its inability to address large positive and negative shocks noticeably differently. Unfortunately, including the independent variance risk premium does not solve this issue (see Figure 3). Nonetheless, it helps mitigate the problem by allowing adjustments to the slope of the NIC. Aside from the initial asymmetry introduced by the shift, these adjustments affect both sides equally, increasing in magnitude with the size of the return shock. Furthermore, a negative independent variance premium universally lifts the level of conditional variance and the lower the premium, i.e., the higher the premium parameter ξ, the steeper the slopes and therefore the reaction to shocks. The special case with $\xi = 0$, yielding the HN model, is displayed as well. An economically rather difficult to justify situation of a positive variance premium would not invert the NIC, but instead lower it and gradually flatten it out.

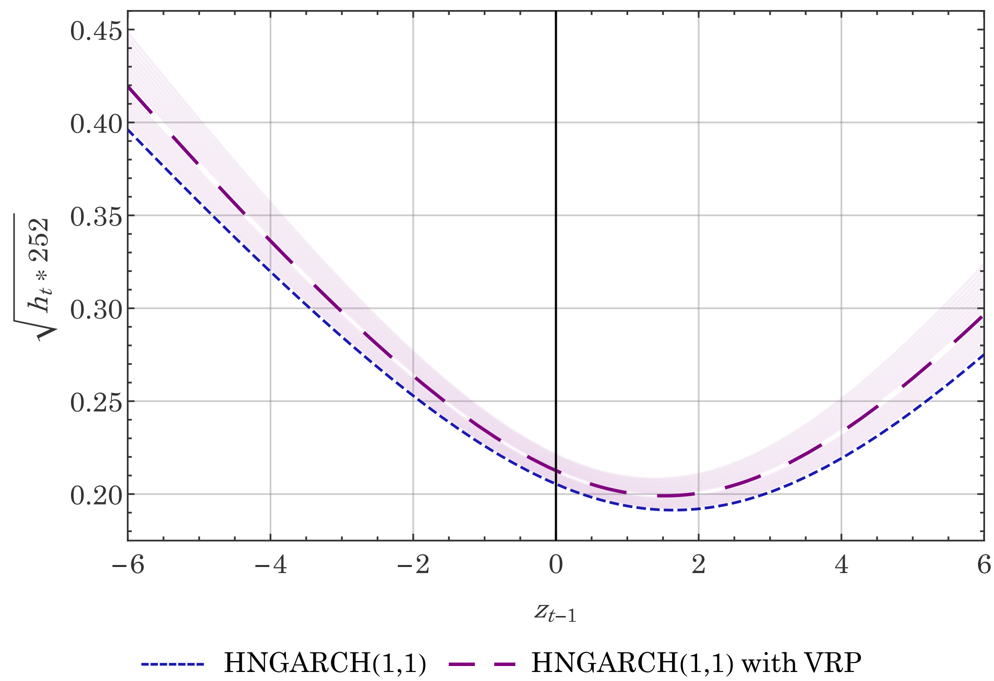

Figure 3 – News impact curve for the Christoffersen et al. (2013) model implied by the parameters in column 3 of Table 5, with different values for the independent variance risk premium parameter. $\xi = 0$ again for the HN case and the bolder dashed curve represents $\xi = 4637$, which is equal to the result obtained in the in-sample estimate.

3 Empirical Analysis

Beginning with a description of the option and return data, the empirical analysis then focuses on the performance of the previously introduced models relative to each other. For this purpose, several scenarios are analyzed and a number of model parameters are tested for their relevance. Furthermore, the importance of the objective function, the method of estimation and the numerical procedures are described. While the BSM model is an obvious benchmark candidate for any option pricing model, we additionally compare the HNGARCH model with the Heston model. Allowing for skewness in the return distribution the Heston model's performance should be much harder to match than the baseline case of the BSM model. Lastly, the extension proposed by CHJ (2013) is included in the comparison and all models are investigated thoroughly, conducting in-sample estimations as well as out-of-sample forecasts, using market data.

3.1 Description of Data

The return sample used in the analysis is taken from the German DAX performance index, i.e., total return index, which is inclusive of dividend payments. Heston and Nandi (2000), like most major empirical studies, rely on S&P500 data, arguing that this index provides sufficient liquidity. However, more than two decades of additional data are available by now and due to increased trading activity and volume, liquidity is not much of an issue anymore. This is especially the case for major indices and common derivatives thereon. Aside, market dynamics have changed rather dramatically compared to the early nineties. Since these changes are also reflected in some fundamental characteristics, conducting an updated evaluation of the model seems like a reasonable intention. As is highlighted in Table 1 the original HN sample exhibits only about a third of the annualized mean return compared to the data used in this work. Further, volatility is less than half in annualized terms, negative skewness is not as prominent and kurtosis is lower as well in their sample, all indicating a much calmer market environment.

While DAX spot prices are freely available from a number of providers they are taken from Thomson Reuters Datastream, which serves us as a credible source. Besides, they all yield virtually identical data, with the exception of an occasional holiday omitted or included. The obtained end of day data points range from May 2nd, 2009 until June 30th, 2016, giving us a total of 1869 daily index returns. Such a long time-series is desirable and necessary for the following estimations. As Christoffersen and Jacobs (2004b), who use 12 years of data, note, it is difficult to obtain precise GARCH parameter estimates from short time-series. Christoffersen et al. (2006), for instance, use roughly 13 years of returns and the data in Hentschel (1995) ranges from the early 1920s until late 1990, including about 64 years of daily observations. With the sample in use, the entire post financial crisis period is covered, giving highly current observations and results. While an even longer time-series would be desirable, one would have to almost double it to somewhat smoothen out the effects of the big plunge in late 2008.

As can be seen in the first graph of Figure 4, the DAX index followed a fairly consistent and steady upward trend. A major plunge only occurred in August 2011 and smaller declines only materialized towards the end of the sample. Over the entire 6-year horizon the index roughly doubled and never dropped by more than 25% from a previous high. The second graph, displaying the corresponding log-returns, provides evidence for volatility clustering. High volatility days are often followed by high volatility days, while periods of relative calmness are rarely disrupted. Clearly visible is the half year of increased volatility following the abrupt drop in mid-2011. Inspired by Barone-Adesi et al. (2008), the last graph in Figure 4 shows the innovations filtered from historical data using the parameter estimates for the HN model from Section 3.2. With a mean of 0.0765 and a standard deviation of 0.9546 these are remarkably close to being standard normally distributed, indicating that our parameters are well identified.

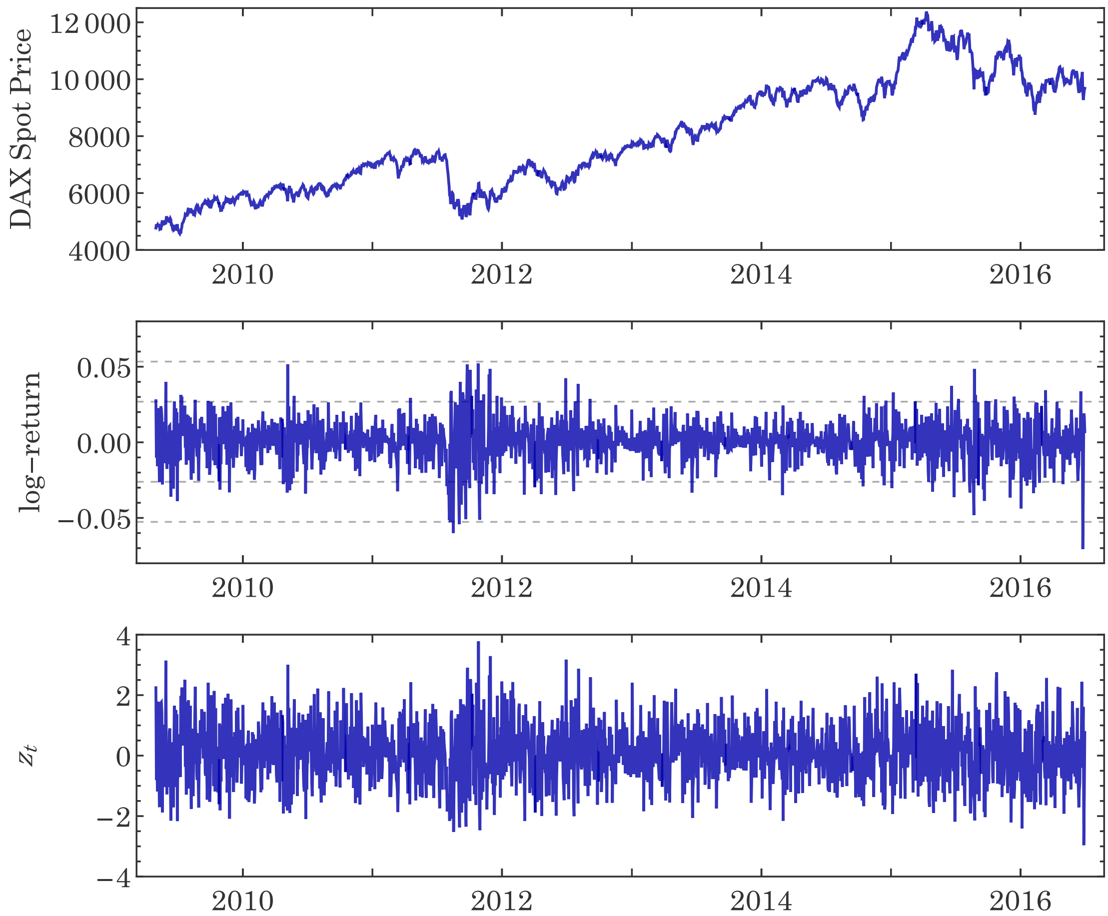

Figure 4 – DAX prices, log-returns and filtered innovations from 1869 observations, ranging from May 2nd, 2009 until June 30th, 2016. Log returns are given with dashed indications for two and four standard deviations from the sample mean.

Just like the returns, option data are obtained from Thomson Reuters Datastream with prices determined in the daily closing auction. These are encompassed by the return sample and they are highly current as well, since they range from early April 2016 until late June 2016. As in Dumas et al. (1998), Heston and Nandi (2000), Christoffersen and Jacobs (2004b) and many other works, the cross-sections are taken each week on Wednesday. Due to computational limitations, this approach allows to still cover a timespan of three months. Furthermore, no Wednesday in our sample is a holiday and studies show that Wednesdays are least affected by volatility or abnormal returns. Giovanis (2009), for example, reports no seasonality in DAX returns for any day of the week and only some seasonality in volatility for Mondays and Tuesdays. Tsiakas (2005) finds highly significant negative returns for Mondays and positive ones on Fridays, while Wednesdays are unaffected. He also reports no significance in monthly returns, except for strong positive effects in September. The cross-sections of option data cover a relatively calm period and a shorter, more volatile period leading up to the "United Kingdom European Union membership referendum", also termed "Brexit referendum". Including the structural break in the data is with the intention to gain a better understanding of how especially the GARCH models adapt to changing market dynamics.

Return and Option Data

Return Characteristics (Annualized)

	Our Sample May 2nd, 2009 – June 30th, 2016	Heston and Nandi (2000) Sample January 09th, 1992 – December 30th, 1994
Number of Observations	1869	755
Mean (%)	10.01	3.23
Standard Deviation (%)	21.04	9.41
Skewness	-0.26	-0.14
Kurtosis	4.99	4.20

Option Data (Number of Contracts)

In – Sample			Moneyness (S/K)			
Time to Maturity (days)	< 0.94	0.94 – 0.98	0.98 – 1.02	1.02 – 1.06	> 1.06	All
< 33	35	43	41	36	43	198
33 – 66	42	49	45	39	52	227
66 – 180	39	49	47	34	46	215
All	116	141	133	109	141	640

Out – of – Sample			Moneyness (S/K)			
Time to Maturity (days)	< 0.94	0.94 – 0.98	0.98 – 1.02	1.02 – 1.06	> 1.06	All
< 33	81	108	94	91	107	481
33 – 66	73	84	73	75	88	393
66 – 180	69	99	84	69	84	405
All	223	291	251	235	279	1279

Table 1 – Return characteristics for the sample used in this work and the sample used by Heston and Nandi (2000). The number of option contracts for the two subsamples is given for each moneyness and time to maturity combination, with the in-sample period being four weeks long and the out-of-sample period comprising data for the following eight weeks.

For this purpose, OTM options with strike prices up to 10% from the respective spot price are obtained on each Wednesday. Given moneyness, defined here as $\frac{S_t}{K}$, smaller than 1 call options are used, while for $\frac{S_t}{K} > 1$ put options are included.[6] These are converted through put-call parity (see equation 2.8), with the interest rate r set accordingly for each cross-section, to avoid immediate arbitrage. Choosing OTM options leads to less exclusions due to a possible lack of liquidity, since a minimum of 5 trades per contract on the observed day is established. Barone-Adesi et al. (2008) note that since the October 1987 crash OTM options have volumes several times higher than in-the-money (ITM) contracts. This is in line with the raw dataset obtained and explained by the desire of investors to insure their positions against larger, especially downside, price movements. Furthermore, options with less than 7 days or more than 180 days until maturity are excluded and only options with monthly expirations are included. The remaining contracts are checked for no-arbitrage conditions along Merton (1973) to guarantees that no profits in excess of r can be made without adequate exposure to risk. Therefore, they must satisfy

$$[S_t - K \exp\left(-r\left(T-t\right)\right)]^+ \leq C_t. \tag{3.1}$$

After applying all criteria, we are left with a total of 1919 options, of which 640 contracts are used for the four week long in-sample period. The following eight week out-of-sample term comprises 1279 contracts. Reasonably dividing them into fairly equally sized categories leads to three maturity and five moneyness groups. Namely, short (<33 days), medium (33–66 days) and long (66–180 days) time to maturity (TTM) and in, at and out, as well as deep in and out-of-the-money (see Table 1).

One could argue that the cross-sections chosen for the empirical analysis are affected by dividend payments, which are usually timed around these months. However, as explained in the theoretical section, dividends are already incorporated in the price of the underlying. Additionally, not all constituents distribute dividends and some of the largest and most heavily weighted companies, such as Siemens or Volkswagen, issue their payments way ahead or after the analyzed period.

Lastly, we could use the prices of the DAX future instead of the cash index levels to infer the required parameters. However, Heston and Nandi (2000) report very similar estimates and virtually unchanged results when working with futures. A possible alternative to calculating with closing prices is proposed by Ting (2006) and applied by Moyaert and Petitjean (2011). Instead of relying on end of day prices he suggests using the volume-weighted average price, short VWAP. Despite appealing findings, this approach is not pursued nor included in our comparison, due to a lack of data.

[6] There exists no standardized definition of moneyness. Heston and Nandi (2000), for instance, define moneyness as $\frac{K}{S_t}$, while Christoffersen et al. (2013) use $\frac{S_t}{K}$.

3.2 Parameter Estimation on Historical Returns

A major advantage of GARCH models over their stochastic volatility counterpart is the ability to straightforwardly incorporate historical information of the underlying. Therefore, due to their inherent structure, using only a single day of data would contradict the fundamental idea of a GARCH model, since the information embedded in past returns would be neglected. Given a sufficiently long time-series, this information can be extracted with a maximum likelihood estimation (MLE). The number of observations is relevant because a MLE is not only consistent, but also asymptotically normal with $\hat{\Theta}_{ML} \overset{a}{\sim} \mathcal{N}\left(\Theta, \frac{1}{I\left(\hat{\Theta}_{ML}\right)}\right)$, where Θ is the true set of parameters, $\hat{\Theta}_{ML}$ its estimator and $I(\cdot)$ the Fisher information. This approach can be seen as a systematic search for the parameter set which maximizes the probability of observing the sample data. More observations should hence provide a better fit to the conditional density of returns and lead to lower standard errors. For both, the HN model and the CHJ (2013) model, the log-likelihood function is equal, since the independent variance risk premium parameter ξ only affects the risk-neutral distribution. It is given by

$$\mathcal{L}^{Returns}\left(\hat{\Theta}_{ML}\right) \propto -\frac{1}{2}\sum_{t=1}^{T}\left(\log\left(h_t\right) + z_t^2\right), \tag{3.2}$$

with $\hat{\Theta}_{ML} = (\alpha, \beta, \gamma, \omega, \mu)$, the five parameters that need to be estimated. A full derivation of the log-likelihood formula can be found in Appendix E. For h_t an initial value needs to be set. Unlike Heston and Nandi (2000), we do not take the variance of the entire sample as a starting value, but only the variance of the previous 15 trading days. This has a number of reasons. Since the return sample is rather long, iterating all the way back to the starting value at every observation is computationally costly. Therefore, at each point the seven preceding values are considered, with the seventh not relying on the eight, but instead on the variance of the 15 previous days. This way 22 trading days of historical information are considered at every step of the optimization routine, which corresponds to roughly one month of time. The log-likelihood did in fact increase with every previous period added, however, less so at each step.

While it appears to be a restrictive assumption and the concept of likelihood does not allow us to compare likelihood values for different sets of initial observations, we can check if the obtained results are economically reasonable and consistent with those of previous works. As it turns out, results are very much in line with other empirical studies. A possible explanation is that the initial value implied by the 15-day variance is usually much closer to the level of conditional variance than the sample variance, therefore requiring less periods to converge. Furthermore, unreported results show that the starting value of the conditional variance, h_0, exerts a negligible influence on the results. As Heston and Nandi (2000) note, this is caused by the strong mean-reversion property of variance.

Inclusive of the one month of initialization in our procedure, the sample for the MLE effectively still ranges from May 2nd, 2009 until June 30th, 2016. The results for various cases are given in Table 2 and are estimated under certain parameter restrictions. α and ω are both required to be larger than zero.[7] β is also limited to non-negative numbers but additionally an upper limit is imposed on the parameter. Reconsidering the formula giving the variance persistence immediately shows why. Even in extreme cases $\alpha\gamma^2$ can never be below zero, irrespective of the sign of γ. To avoid an integrated variance process, the sum in equation (2.27) and therefore β has to be smaller than one. Aside from the special case of $\gamma = 0$, the skewness parameter is estimated freely. However, unless extremely short samples with temporary positive skewness are used, it always remains positive, indicating predominantly negative skewness in daily DAX returns.

Estimating μ without any restriction leads to widely fluctuating values, even reaching negative territory at some points. Giving it a reasonable range of starting values avoids this issue without imposing any limitations. Moreover, as CHJ (2013) note, realistic values for μ can quickly be inferred from market data. An annual equity risk premium given by $\mu h_t = 10.01\%$ and an annualized volatility of about 21.04% implies $\mu = 2.26$. These are approximate average return and volatility values for the years of the sample in use (see Table 1). Looking at longer time horizons with an average of roughly $\mu h_t = 8\%$ and $\sqrt{h_t} = 18\%$ yields $\mu = 2.47$. Both remarkably close to the estimated value in Table 2 of $\mu = 2.49$, again confirming the plausibility of our results.

Motivated by the original work of Heston and Nandi (2000), we analyze the case of $\gamma = 0$ to determine the importance of the skewness parameter. Even though this restricted estimation leads to substantial adjustments in the remaining parameters they are not sufficient to obtain an equally high log-likelihood. In an attempt to capture what is else described by the skewness parameter, α, β and μ increase. Note especially the changed scale of α and the virtually identical persistence, despite different estimates. Since $\alpha\gamma^2 = 0$ now, the entire persistence is explained by the parameter giving weight to previous conditional variances, β. This also signals the relevance of the persistence, which seems to be a priority in estimation. All of our findings, including the decrease of ω, are in line with Christoffersen and Jacobs (2004b). Christoffersen and Jacobs (2004b) also provide a possible explanation to why the parameter responsible for creating, i.e., describing, kurtosis in the distribution, α, increases when γ is reduced or eliminated. Since both control the value of options with strikes far from the spot price they might, to a certain extent, be substitutable. Furthermore, as long as $\gamma \neq 0$, this also takes pressure from β by keeping persistence in line. Similar results, of a high negative correlation between α and γ, were observed as well when extracting the parameter values on a daily basis from our estimation.

[7] Technically, as long as α is positive, ω is allowed to reach zero in the HN model and usually small values are estimated. For our estimation this restriction was not binding, however.

To gain a more intuitive understanding of these persistence values, we calculate the half-life of a shock according to Nelson (1991). The simple ratio, $\mathcal{H} = \frac{\log(0.5)}{\log(Persistence)}$, gives the number of days it takes for a return shock to fade by half. For all three scenarios this is roughly nine trading days, with shocks under a higher persistence consequently taking longer to fade.

Maximum Likelihood Estimation

Physical Parameters	No Skewness (γ=0)		No Premium (μ=0)		HN(2000) Model	
α	1.02E-05	(1.18E-06)	8.17E-06	(3.03E-07)	8.17E-06	(3.43E-07)
ß	0.926	(0.006)	0.806	(0.002)	0.806	(0.006)
γ	–	–	124.01	(3.10)	121.56	(3.10)
ω	1.45E-06	(8.85E-07)	3.75E-06	(7.52E-07)	3.76E-06	(8.06E-07)
μ	5.070	(1.773)	–	–	2.491	(1.523)
Long–Run Volatility	19.91%		21.02%		20.29%	
Persistence	0.9257		0.9320		0.9270	
Half–Life (days)	8.98		9.84		9.14	
Log–Likelihood	7211.36		7254.38		7255.35	
p–Value	0.0000		0.1626		–	

Table 2 – Parameters of the maximum likelihood estimation for the restricted cases of zero skewness ($\gamma = 0$), zero equity premium ($\mu = 0$) and for the unrestricted HN (2000) model. Corresponding standard errors appear in parentheses. The annualized long-run volatility is defined to be equal to $\Omega = \sqrt{\frac{252(\omega+\alpha)}{1-\beta-\alpha\gamma^2}}$ and persistence of variance is measured by $\beta + \alpha\gamma^2$. Half-life, indicating the time in days it takes for shocks to fade by half, is given as $\mathcal{H} = \frac{\log(0.5)}{\log(Persistence)}$ and the p-value stems from the likelihood-ratio test statistic, $LR = -2\,(LL_R - LL_U) \sim \chi^2\,(k)$.

The annualized long-run volatility is easily derived from equation (D.7) and is given by

$$\Omega = \sqrt{\frac{252\,(\omega + \alpha)}{1 - \beta - \alpha\gamma^2}}. \tag{3.3}$$

Since the variance mean-reversion in the denominator is virtually identical for the symmetric and asymmetric case in Table 2, the difference in long-run volatility mainly depends on the sum of α and ω. The highest value for parameter Ω is observed for the zero premium case and is caused by the slightly higher skewness parameter.

CHJ (2013) apparently impose $\mu = 0$ after having obtained the original HN model parameters in their sequential estimation. They do not provide an explanation for this choice, which is rather peculiar since they estimated the model conventionally when pursuing a joint approach. We estimate the model again for our data, given the zero asset premium, therefore allowing adjustments in the other parameters, which are, however, barely noticeable. The resulting impact on return log-likelihood is very limited, irrespective of allowing adjustments or not. To test for significance, we conduct a likelihood ratio test. This is a simple, chi-squared distributed test statistic which allows us to compare differently nested models among each other and also to the full, unrestricted model. The ratio is given by

$$LR = -2\left(LL_R - LL_U\right) \sim \chi^2\left(k\right), \tag{3.4}$$

with LL_R and LL_U the log-likelihood values of the restricted and unrestricted model, respectively. k stands for the degrees of freedom and is equal to the number of omitted parameters. If one or more parameters are insignificant their elimination should not lead to a large reduction in the maximized log-likelihood. The resulting p-values are given in the last row of Table 2. As expected γ is a highly relevant parameter and should thus be included in the estimation. The p-value of $\mu = 0$ on the other hand, even though not extremely far above 10%, does not lead to a significant reduction in log-likelihood using commonly applied levels. This can be attributed to its relatively high standard error, which indicates a rather flat log-likelihood function in this dimension. Moreover, γ increases by about as much as μ is reduced, likely absorbing most of its effects.

Asymptotic standard errors, which are displayed in parentheses, are given by the inverse of the square root of the diagonal elements contained in the Fisher information matrix. The Fisher information is the negative of the expectation of the Hessian, i.e., the second derivative, of the log-likelihood function. Therefore, the standard errors can be calculated according to

$$S.E.\left(\hat{\Theta}_{ML}\right) = \frac{1}{\sqrt{I\left(\hat{\Theta}_{ML}\right)}}, \tag{3.5}$$

with

$$I\left(\hat{\Theta}_{ML}\right) = -\mathbb{E}\left[\left.\frac{\partial^2 \mathcal{L}\left(\Theta\right)}{\partial\Theta_i \partial\Theta_j}\right|_{\Theta=\Theta^*}\right], \tag{3.6}$$

where $1 \leq i, j, \leq p$ and p the number of parameters. While no clear pattern can be identified between the three cases, regarding the relative size of the standard errors, we can group the parameters themselves. As already noted μ has a rather large standard error, which is almost two thirds of the estimated value in the unrestricted case. To a lesser extent, the same holds for ω. The remaining parameters, α, γ and especially β, have relative standard errors as low as $\frac{0.006}{0.806} = 0.74\%$. It is therefore crucial to estimate these parameters accurately and precisely, since any deviation has comparably large implications. Contrarily, less emphasis should be put on estimating ω or μ. Aside from slightly lower standard errors, setting $\mu = 0$ has no significant implications when fitting to returns. In the final result, a higher or lower parameter, even in the order of magnitude of 50%, is barely noticeable.

Some authors, such as Byun (2011) and CHJ (2013), obtain slightly negative or very small values for the parameter ω and consequently set it equal to zero. While this lowers the computation time it leads to significantly lower log-likelihoods in our case. We therefore estimate ω freely and provide the alternative suggestion of setting it to a reasonably small number in the region of 1×10^{-7} to 1×10^{-6}. This range is supported by results of a large number of reputable works (e.g., Heston and Nandi (2000), Christoffersen and Jacobs (2004b) and Christoffersen et al. (2012)), as well as our own results and explains between 1.6% and 5% of annual conditional standard deviation. Combining this insight with the suggestion to infer μ from market data, not just to check for plausibility, but to omit it from estimation, reduces the calibration problem from five to three parameters. Taking a look at equation (2.26) we can see that ω functions as a lower bound, or base level, for the conditional variance. Assuming conditional volatility for DAX returns does not reach levels that low, including ω leaves less explaining to the remaining parameters, possibly allowing them to fit the actual structure or certain patterns better.

Furthermore, we can use equation (2.28) and calculate the conditional variance for each day of our sample, by simply plugging in the parameter values from the last column of Table 2. Following Heston and Nandi (2000) and doing the same for the case of zero skewness, once again highlights the importance of the asymmetry introduced by γ (see Figure 5 for an annualized representation). Compared to the unrestricted model, the symmetric model exhibits a much less volatile process over time. Not only does it lack sudden drops in the filtered conditional variance, it also trails the unrestricted model in size and suddenness of increases. Moreover, it generally seems to deviate less from its long-run mean and also return after fewer periods, what is in line with the lower persistence value.

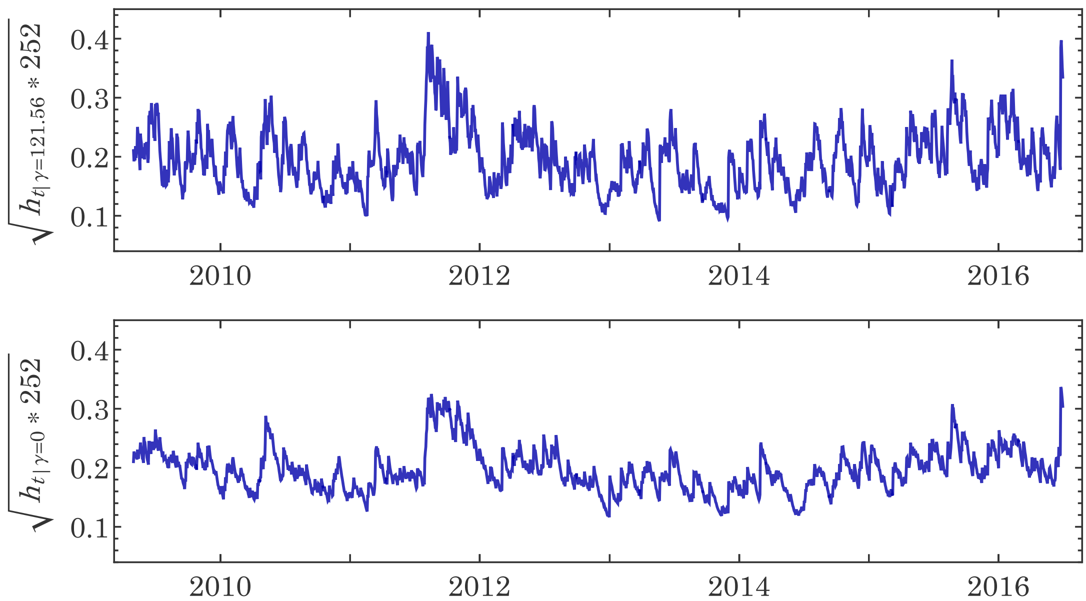

Figure 5 – Annualized conditional volatility for the asymmetric and symmetric GARCH model for 1869 observations, ranging from May 2nd, 2009 until June 30th, 2016.

With the available parameters we can use the covariance equation (2.31) and the variance of conditional variance in equation (D.12) to calculate the correlation between returns and future conditional variance. This leads to

$$Corr\left(R_t, h_{t+1}\right) = \frac{-2\alpha\gamma h_t}{\sqrt{2\alpha^2\left(1+2\gamma^2 h_t\right)\times h_t}}. \tag{3.7}$$

Given h_t is at its long-run value, which we set equal to the sample variance, yields an annualized volatility of $\sqrt{h_t \times 252} = 21.04\%$ and a correlation of -0.9157. As Heston and Nandi (2000) describe, such a near perfect negative correlation should lead to a situation where variance behaves like a negative beta asset.

For the MLE, but also for the following optimizations, it is important to employ numerical methodologies that can cope with the complex structure of the models. In most situations the Nelder and Mead (1965) direct search algorithm is applied, since it proves to be a well-balanced choice between accuracy and cost of computation. However, for the few cases in which different sets of starting values lead to inconsistent results, we rely on the slower but much more precise "Differential Evolution" technique as proposed by Storn and Price (1997). While both of these methods still don't guarantee a true global minimum they are able to skip local minima. Computation tends to take longer but ultimately the results are better, since the distance between observations and model values is smaller. We follow the suggestion of Gilli and Schumann (2012) and estimate each case multiple times with varying starting values. This allows us to get an idea of the sensitivity of our results and also avoids conjecturing based on a single outcome. Further, as they note and as we've seen, especially for the asset premium μ, different sets of parameters may also lead to satisfactory results. However, they imply different dynamics and since we intend to use them throughout the analysis to value options, this is not desirable.

3.3 Model Calibration Using Option Data

Having obtained the physical parameters for the GARCH models from the return series, we now proceed to value options. Following the sequential estimation approach of CHJ (2013) we take advantage of the knowledge of mapping and directly obtain the required risk-neutral skewness parameter for the HN model. For the more sophisticated model with augmented pricing kernel we still need to estimate the remaining independent variance risk premium parameter using option data in order to complete the mapping. Equally, the BSM model has one unknown parameter, the implied volatility, which needs to be determined. A more complex case is the calibration of the Heston model. In addition to its four time-invariant structural parameters we have an initial variance for each cross-section of option data. While estimating all these parameters seems like a straightforward procedure, there are a few considerations that have to be taken into account.

3.3.1 Choice and Importance of the Objective Function

Firstly, and most importantly, it is essential to define the objective function which is used to calibrate the models. A popular approach is to minimize, one way or another, the error between model prices and market data. However, depending on the choice of objective function, varying weights are implicitly assigned to the option contracts. These then lead to different parameter estimates and ultimately different results. Therefore, it is essential to apply the same objective function to all of the models and to rely, ideally, on more than just one loss function for comparison. Nonetheless, even with the same objective function, it is still difficult to compare continuous and discrete-time models.

Furthermore, due to computational limitations we obtain the risk-neutral parameters for the HN model solely from returns and corresponding mapping instead of conducting a joint estimation. While we can still evaluate the loss function using these parameters they might be suboptimal. That is, there exists a parameter set, obtainable through joint estimation, which achieves a better fit to option data in exchange for a lower return log-likelihood. Ultimately, even a large joint estimation, as conducted in CHJ (2013), faces the issue of a trade-off between option and return fit, with the importance assigned to each depending on the ratio of observations. Since the option data usually carries much more weight, even when the time-series of returns is long, the supposedly joint estimation effectively turns into a two-step process (Bates, 2003). Risk-neutral parameters are estimated first and the remaining physical parameters are fitted afterwards, which greatly affects standard errors. This is also noticeable in CHJ (2013), where the risk-neutral parameters obtained in a joint estimation only differ by a margin from a purely option based optimization, despite a relatively long return sample. Since all options in our analysis are affected by the same underlying it makes sense for the GARCH models to share the same time-series parameters. In our sequential estimation we get these through MLE (see Table 2).

Proceeding to measures that allow us to compare the resulting model errors, it is obvious that simply minimizing the total error between option and market prices is not sufficient. Since negative and positive deviations are possible, these are not necessarily reduced but simply matched by an optimization routine. A common solution is to use absolute values, or to square the errors and report the square root of their mean. This can either be done with prices, but also with relative prices, or even implied volatilities. When minimizing the pricing error, we obtain the euro root mean squared error (€RMSE), which is given by

$$€RMSE\left(\Theta\right) = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(C_i - C_i\left(\Theta\right)\right)^2}, \tag{3.8}$$

with N representing the number of option contracts, C_i the market price of option i and $C_i\left(\Theta\right)$ the model price for that same option, calculated with the set of parameters, Θ. This loss function particularly emphasizes options with higher prices, i.e., ITM call options with long TTM in our case, since these cause potentially larger errors. Despite this known bias, it is still widely used in, for instance, works of Bakshi et al. (1997) or Heston and Nandi (2000).

Another objective function, which is utilized by Christoffersen et al. (2010b), is the relative root mean squared error (RRMSE). It is quite similar to the €RMSE, however, the errors are divided by the analogous market price before squaring them. With C_i again as the market price, it is given by

$$RRMSE\left(\Theta\right) = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\frac{C_i - C_i\left(\Theta\right)}{C_i}\right)^2}. \tag{3.9}$$

Everything else is virtually identical, but given the division, this approach favors options with small values and therefore deep OTM calls. Generally, this leads to parameters attempting to fit options without intrinsic value and relatively short maturities, which are likely to expire worthless. Since these contracts tend to face liquidity issues data can be very noisy and possibly unreliable. Therefore, despite its intuitive interpretation, this objective function is rarely used in option pricing literature.

A more balanced choice of objective function, lending similar weight to all options, is the implied volatility root mean squared error (IVRMSE). Irrespective of the moneyness or time to maturity of a contract, implied volatility is always, more or less, of similar size. Minimizing the difference between implied volatilities leads to

$$IVRMSE\left(\Theta\right) = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\sigma_i - \sigma_i\left(\Theta\right)\right)^2}, \tag{3.10}$$

where σ_i represents the implied volatility for the market price of option i and $\sigma_i\left(\Theta\right)$ the model implied volatility. These are obtained by inverting the BSM formula, setting it equal to the respective price and solving numerically for the volatility parameter σ.

We will estimate the physical parameters of the models using the €RMSE and report and compare the resulting pricing errors. Further, we will separately estimate the models using equation (3.10) as the loss function and report the IVRMSE. This is different from, for instance, Schoutens et al. (2003), who obtain model parameters by minimizing the squared pricing errors and then plugging these into all of their loss functions. While our selection seems arbitrary, it is based on the popularity of the objective functions. For practical applications the choice of objective function should depend on the purpose of the model and, as Christoffersen and Jacobs (2004a) note, on the end-use of the model.

3.3.2 In-Sample Estimation

The in-sample estimation, which we use to calibrate the models, is conducted on the first four cross-sections of our option data. Following the same order in which we've introduced the models, we begin by estimating the only parameter of the BSM model, σ. Since the model is allowed a separate parameter for each week anyways, there is no real need to calibrate it on multi-week data. However, we still report the values for comparison and obtain these by minimizing the loss functions in equations (3.8) and (3.10). The optimized parameters are given in Table 3 and differ only slightly at first glance. While for the €RMSE we actually have to minimize the price differences and find suitable parameters, for the IVRMSE loss function we can simply take the mean of the market implied volatility each week. By definition this minimizes the error in the BSM model.

σ_{BS}	Week 1	Week 2	Week 3	Week 4
€RMSE	0.2323	0.2070	0.2081	0.2105
IVRMSE	0.2392	0.2016	0.2109	0.2132

Table 3 – In-sample parameter estimates for the BSM model under both objective functions.

Unlike the BSM model the Heston model has a set of time-invariant parameters which are held constant over all four weeks. The initial variance, however, is estimated for each cross-section, leaving us with a total of eight unknown values. Table 4 shows the results of the calibration. Again, the estimates are fairly similar across the two objective functions. While θ^* is a bit higher when minimizing the IVRMSE, the initial variances are all lower in exchange. This seems, just like the previously mentioned tradeoff between kurtosis and skewness very plausible. In the Heston model these two moments are represented by ϱ, which is higher for the IVRMSE and ρ, which is also higher, therefore indicating a lower negative skewness. Lastly, κ^* signals a slightly lower mean-reversion speed.

€RMSE				IVRMSE			
ϱ	1.0214	v1	0.0683	ϱ	1.2851	v1	0.0631
ρ	−0.6948	v2	0.0451	ρ	−0.6772	v2	0.0392
κ^*	3.7500	v3	0.0464	κ^*	3.4997	v3	0.0450
θ^*	0.0660	v4	0.0483	θ^*	0.0815	v4	0.0471

Table 4 – Heston model parameter estimates for the two objective functions.

Since we're estimating the model only on option data we yield a risk-neutral set of parameters. These do not allow us to identify the variance risk premium λ, which is inseparably included in $\kappa^* = \kappa + \lambda$ and $\theta^* = \frac{\kappa\theta}{\kappa+\lambda}$. As long as we're only attempting to value contingent claims this is not a problem, because we just need the aggregate values. However, compared to the maximum likelihood approach using returns, where we can independently estimate the parameters and combine them as needed, we lack some information.

Another weakness of the purely option based estimation approach results from the initial variance. For each additional cross-section, we have to estimate another value. This quickly leads to a highly parameterized problem as Christoffersen et al. (2010b) note and aside from that, it also takes longer and longer to estimate the model. A possible way to speed up calibration is suggested by, for instance, Kilin (2011). Reducing the number of computationally costly evaluations of the characteristic function by caching its values avoids unnecessary recalculations. As can be seen from equation (2.17) the CF does not depend on the strike price K therefore it only needs to be computed once for each parameter and TTM combination. Furthermore, integration time can be reduced greatly by relying on alternative integrations schemes, such as Gaussian quadratures. Due to the semi-infinite integration range of our problem the Gauss-Laguerre quadrature is of particular interest. Since the integrands converge rapidly it is sufficient to only evaluate them at relatively few nodes when applying this method, without producing noticeable errors (Schmelzle, 2010, P. 44). Alternatively, estimating all parameters can be avoided by presuming some of the values. The state variable ν_t can be approximated by a short term at-the-money (ATM) implied volatility. Another possible suggestion by Clark (2011, P. 130–132) sets the mean-reversion rate when pricing foreign exchange options to $\kappa = \frac{2.75}{T_{cal}}$ and $\theta = \sigma^{BS}_{ATM}(T_{cal})$, where T_{cal} is the calibration horizon.

Of the GARCH models, only the CHJ (2013) model needs additional estimation. For the HN (2000) model we rely on the mapping of the skewness parameter and recall that $\gamma^* = \gamma + \mu$. Simply summing up the physical skewness parameter and the asset risk premium allows us to quickly calculate the value for risk-neutral skewness. For the special case with zero premia the risk-neutral parameters are all identical to those obtained in the MLE and do not provide any specific fit to the in-sample data. Together with an option based estimate of the independent variance risk premium ξ, the resulting risk-neutral parameters and properties are given in Table 5. Regarding ξ our estimation approach differs somewhat from CHJ (2013) in that we use the same objective functions as for the BSM model and the model of Heston. While we minimize the squared pricing errors, CHJ (2013) additionally divide each error by the Black-Scholes vega of the respective option at its market implied volatility level. Our tests show that the resulting estimates do not differ by much, therefore this can be seen as a suitable and fast approximation. However, since our sample is much smaller and the only remaining parameter is identified quickly, we do not need to rely on approximations. Again, due to the sequential nature of our estimation approach, we can take advantage of the possibility of mapping and use equations (2.44)–(2.46). This yields the remaining risk-neutral parameters and completes the in-sample calibration of our models.

Risk – Neutral Parameters and Properties

	No Premia (μ=ξ=0)	HN (2000) Model	CHJ (2013) Model	
			€RMSE	IVRMSE
ξ	–	–	4637	6433
1/(1-2αξ)	1	1	1.0820	1.1174
α*	8.17E-06	8.17E-06	9.56E-06	1.02E-05
β*	0.806	0.806	0.806	0.806
γ*	124.01	124.05	114.69	111.06
ω*	3.75E-06	3.76E-06	4.06E-06	4.20E-06
Long–Run Volatility	21.02%	21.02%	22.48%	23.12%
Persistence	93.20%	93.20%	93.21%	93.21%
Half–Life (days)	9.84	9.84	9.85	9.86
Correlation (R(t),h*(t+1))	-0.9186	-0.9186	-0.9067	-0.9014
Total Likelihood	4591.40	4594.47	4757.67	4719.79
From Returns	7254.38	7255.35	7255.35	7255.35
From Options	-2662.98	-2660.88	-2497.68	-2535.56

Table 5 – Risk-neutral parameters for the CHJ (2013) model, as well as the nested models of HN (2000) for $\xi = 0$ and the special case without any risk premia. The values in the first two columns are identical to the physical parameter estimates, given $\gamma^* = \gamma + \mu$. The two columns to the right contain the values for a freely estimated independent variance risk premium parameter under the two different objective functions used in calibration.

The risk-neutral parameters reported in the first column of Table 5 are identical to the physical parameters. This is easily explained by the absence of any risk premia. For the HN model we yield a risk-neutral skewness parameter almost identical to the no premia case. For both situations the scaling factor $1/(1-2\alpha\xi)$ is equal to 1 because the independent variance risk premium is assumed to be zero. A nonzero estimate of ξ leads to a further refinement of the wedge between the empirical and risk-neutral distributions. Just like in column three and four, for $\xi > 0$ the variance risk premium is negative and increases the scaling factor, which in return affects all parameters except for β, ultimately putting more weight on the tails of innovations (Christoffersen et al., 2010a). While the parameter estimate itself is almost 50% higher in the IVRMSE case the scaling factors for the two objective functions only differ by roughly 3%, leading to similar parameters.

We can also see that the estimate of ξ is closely related to α. To keep the ratio between the two probability measures in line, a decrease of α tends to lead to an increase in ξ. Furthermore, with knowledge of α we can determine an upper limit for the independent variance premium parameter. Once $2\alpha\xi = 1$ the scaling factor implodes and even turns negative if ξ is increased further. While this would still allow us to obtain plausible option prices, the resulting properties from such a parameter combination are very different. Therefore, one cannot simply minimize the difference between model and market price but has to take into account the economic implications. Imposing the restriction $\xi < \frac{1}{2\alpha}$ avoids obtaining a troubling optimum. Given parameter $\alpha = 8.1688 \times 10^{-6}$, as estimated from the time-series of returns, this implies $\xi \overset{!}{<} 61208.7$.

Compared to the case of a monotonically declining pricing kernel in the HN model, all affected parameters are notably different for the CHJ (2013) estimations. ω^* is higher, since ω is multiplied by the scaling factor and α^* deviates even more since α is multiplied by the squared scaling factor. The lower value for γ^* might seem counterintuitive at first, however, the reduction is relatively small compared to the increase in the other parameters and as we've noted earlier, skewness and kurtosis offset each other to some degree. These changed parameters lead to an increased wedge and can explain the empirically observed properties of fatter tails and higher kurtosis in options compared to returns (Aït-Sahalia and Lo, 1998; Jackwerth, 2000).

By definition, for a positive return premium ($\mu > 0$), a negative independent variance risk premium ($\xi > 0$) and a negative correlation between returns and variance ($\gamma > 0$) the risk-neutral conditional variance has to exceed the physical conditional variance. With the estimated physical annualized long-run volatility at 20.29%, the HN model sees a rather modest increase when changing measure. Much more pronounced is the difference in risk-neutral long-run volatility for the two estimations taking advantage of the U-shaped pricing kernel. Both scenarios give values which are more than two percentage points higher. Despite the differences in volatility the risk-neutral persistence increases to an almost perfectly equal level for all for estimations. Consequently, the closely related half-life is almost identical as well. This once again confirms the relation between α and γ, or for the risk-neutral case α^* and γ^*. With the risk-neutral persistence given by equation (2.51) and β^* the same across all four columns, the persistence can only stay constant if changes in α^* and γ^* balance out.

The risk-neutral correlation between returns and conditional variance for the following period is highly negative for all four cases and just like for the physical parameters it is based on the sample variance. The obtained likelihood values are calculated with the following log-likelihood formula

$$\mathcal{L}^{Options}\left(\hat{\Theta}^*_{ML}\right) \propto -\frac{1}{2}\sum_{i=1}^{N}\left(\log\left(\hat{s}^2\right) + \frac{\left(C_i - C_i\left(\hat{\Theta}^*\right)\right)^2}{\hat{s}^2}\right), \tag{3.11}$$

where $\hat{s}^2 = \frac{1}{N}\sum_{i=1}^{N}\left(C_i - C_i\left(\hat{\Theta}\right)\right)^2$, i.e., the mean squared error of the sample and N the total number of option contracts. $\hat{\Theta}^*_{ML} = (\alpha^*, \beta^*, \gamma^*, \omega^*)$ is the set of four risk-neutral parameters used to obtain the log-likelihood. With similar parameter values, the HN (2000) barely exceeds the no premia case, again pointing to the limited influence of the equity risk premium. Just like for the Heston model and as is generally the case, the equity risk premium can only be estimated from the time-series of returns. As expected, the CHJ (2013) model yields a much better fit to option data because we estimate ξ accordingly. This leads to highly significant improvements and is also reflected in the total likelihood, which is just the sum of return and option likelihood. While expected, we also see that the implied volatility based parameters lead to a much lower option log-likelihood, supporting the approach of different parameters for each objective function. For the BSM model we get a value of -2718.39, which is the lowest likelihood and for the Heston model a value of -1891.53 when the option prices obtained with their parameter estimates are plugged into equation (3.11).

Additionally, unreported results show that when starting from the sample mean it only takes about 15 periods for the ratio of risk-neutral conditional variances between the HN and the CHJ model to settle at the respective level imposed by the scaling factor. This is another indication that an arbitrary but reasonable initial value only has a very limited effect, which then completely vanishes. Furthermore, due to the constant ratio, especially during volatile times the difference between these two becomes more pronounced in absolute values. Their means over the entire sample length are 19.54% for the HN and 20.32% for the CHJ model, with the CHJ (2013) approach at any time at least 0.37% and at most 1.65% higher for $\xi = 4637$.

Regarding estimation procedure, it can be seen from equation (2.37) that the CF for the GARCH models does also not depend on the strike price K. This allows to apply the previously described methods and speed up calculation, therefore enabling us to process more data. Using more than just one cross-section is preferable since it allows us to capture time variations in the data. Furthermore, the supposedly time-invariant parameters, which should fit this aggregate information, would vary over time if we chose to recalibrate the model on single cross-sections, as is often done by practitioners (Müller et al., 2013).

To compare the model fit we plug the estimated parameters in the respective models and evaluate the objective functions. The aggregate results for the in-sample period are presented in Table 6 and the models are sorted in ascending order of errors. With a RMSE of less than 12€ and an IVRMSE of about one percentage point the Heston model clearly

Aggregate In – Sample Valuation Errors

	€RMSE	%€RMSE	Average Price	IVRMSE	%IVRMSE	Average IV
Heston	11.71	2.40%	488.08	0.94%	4.35%	21.67%
CHJ	30.04	6.16%	488.08	2.23%	10.28%	21.67%
HN	38.77	7.94%	488.08	2.83%	13.05%	21.67%
BSM	42.75	8.76%	488.08	3.50%	16.16%	21.67%

Table 6 – Aggregate in-sample valuation errors for the models sorted in ascending order. The results are given in absolute terms as well as in relation to the average price and average implied volatility of the 640 contracts.

offers the best fit. This is expectable since it is the only model where more than one parameter is actually calibrated to the option data. Setting these errors in relation to the average option prices and average implied volatilities, which can be found in Table 10 in the Appendix, yields percentage deviations from the market prices of 2.4% and from the market implied volatilities of 4.35%. These percentage results are not to be confused with the RRMSE from equation (3.9). They are simply intended to facilitate comparison to other samples of data.

Regarding the two GARCH models the CHJ (2013) and its independent variance premium lead to a reduction of errors as to the HN model of about one fifth. Nonetheless, in total it still produces much larger total errors than the Heston model. Last and worst performing of the models is the BSM model. Despite weekly recalibration of its parameter to the option data the raw pricing RMSE amounts to more than 42€ and a relative IVRMSE of over 16% is clearly worst as well. Considering the comparably stable period over which we conducted the in-sample estimations, our models produce results which vary by quite a bit. It is, however, important to keep in mind that these are purely fitting exercises and do not necessarily reflect the ability of models to forecast option prices. Furthermore, based on estimation technique, the HN model is not calibrated to this set of options and therefore cannot be an optimal solution.

Splitting up the results further and taking a look at the disaggregate data provides a more detailed insight. The values given in Table 7 are the actual errors and are not divided by the category averages. As expected, the Heston model is fitted especially towards the longer-end of our option data when attempting to minimize the €RMSE. Despite ever higher option prices the absolute errors reduce for longer maturities and higher moneyness categories. This gets even more obvious when looking at the third column in Table 7, showing pricing errors for 66–180 days, or at relative values (see Figure 6). The plotting of percentage errors also enables us to see that due to its near perfect fit to high priced options, the Heston model is outperformed at some nodes for shorter maturities.

On the other hand, the return data estimated HN model provides a surprisingly good fit to OTM calls in general. For the shorter maturity category, it even outperforms the HN model at some instances. It will be interesting to see how well the model fares out-of-sample, but for now we see that the calibration on returns leads to parameters which fit short-term OTM calls best. Furthermore, even though the HN model does not offer the best fit, when looking at the percentage plots in Figure 6 we see that it leads to very steady errors. Including the additional flexibility provided by the CHJ model leads to convergence of the different errors and it makes up for its worse performance in lower moneyness groups by clearly outperforming the HN model for ATM and ITM contracts. This becomes especially apparent for longer maturities where the CHJ approach leads to an almost symmetrical curve. Setting these errors in relation to the average option prices again highlights the focus on higher priced options. It also shows that aside from deep OTM Calls the performance,

In – Sample Valuation Errors Disaggregated by Moneyness and Maturity

		€RMSE				IVRMSE			
		< 33	33 – 66	66 – 180	All	< 33	33 – 66	66 – 180	All
< 0.94	Heston	4.59	7.12	13.84	9.44	0.44%	0.76%	0.77%	0.68%
	CHJ	7.26	15.14	51.81	31.64	1.78%	2.00%	2.50%	2.12%
	HN	3.68	6.52	20.05	12.44	1.16%	0.65%	0.74%	0.86%
	BSM	24.55	44.14	62.86	47.07	4.89%	4.25%	2.62%	4.01%
0.94 – 0.98	Heston	11.68	11.12	14.29	12.47	0.98%	0.99%	0.76%	0.91%
	CHJ	13.35	14.63	36.41	24.28	2.11%	1.44%	1.75%	1.77%
	HN	10.01	14.59	18.92	15.13	1.38%	0.85%	0.67%	1.00%
	BSM	31.90	38.00	40.74	37.27	4.02%	2.80%	1.75%	2.95%
0.98 – 1.02	Heston	16.66	11.03	13.35	13.77	1.33%	0.81%	0.77%	0.99%
	CHJ	15.87	16.14	22.48	18.56	1.66%	0.80%	1.07%	1.21%
	HN	20.99	32.44	37.40	31.40	1.84%	1.83%	1.35%	1.68%
	BSM	17.01	15.98	16.53	16.50	1.70%	0.95%	0.88%	1.21%
1.02 – 1.06	Heston	10.45	10.33	10.48	10.42	0.87%	0.57%	0.72%	0.73%
	CHJ	24.13	31.99	29.51	28.81	2.25%	1.55%	0.80%	1.65%
	HN	32.36	49.96	62.52	49.58	3.45%	3.14%	2.35%	3.03%
	BSM	25.32	38.85	43.46	36.62	2.62%	2.22%	1.56%	2.19%
> 1.06	Heston	12.31	12.24	9.42	11.42	1.94%	0.56%	0.83%	1.22%
	CHJ	30.44	45.71	45.12	31.44	5.11%	3.11%	1.44%	3.49%
	HN	36.05	60.15	76.41	60.30	6.52%	4.81%	3.26%	5.00%
	BSM	38.99	63.59	74.99	61.56	7.34%	4.98%	3.08%	5.35%

Table 7 – In-sample valuation errors of the different models disaggregate by objective function, moneyness and TTM.

despite the strong inherent bias from the return estimation, is really good and not too far from the Heston model. The BSM model has its best fit near and ATM, even beating the other models at some nodes. It produces almost equal absolute errors for options with the same distance to the spot price, therefore failing especially for lower moneyness groups when looking at relative errors.

Our results for the IVRMSE are somewhat similar, however, due to the almost equally balanced objective function models calibrated to options exhibit a more symmetrical loss distribution across the moneyness categories. The observed market volatility surface is fairly flat in time ATM, slightly upwards sloping for OTM calls and notably downwards sloping for options with higher moneyness. This results in a more pronounced skew for contracts near expiry, which flattens out gradually towards longer maturities. Again, the Heston model is able to match market data best. Putting less emphasis on options with higher prices also helps to fit short term OTM calls really well now. As we can see in the second column from the right in Table 7, the Heston model produces an almost flat error curve for longer dated contracts. This comes at the price of a weaker and more fluctuating performance for the shortest maturity.

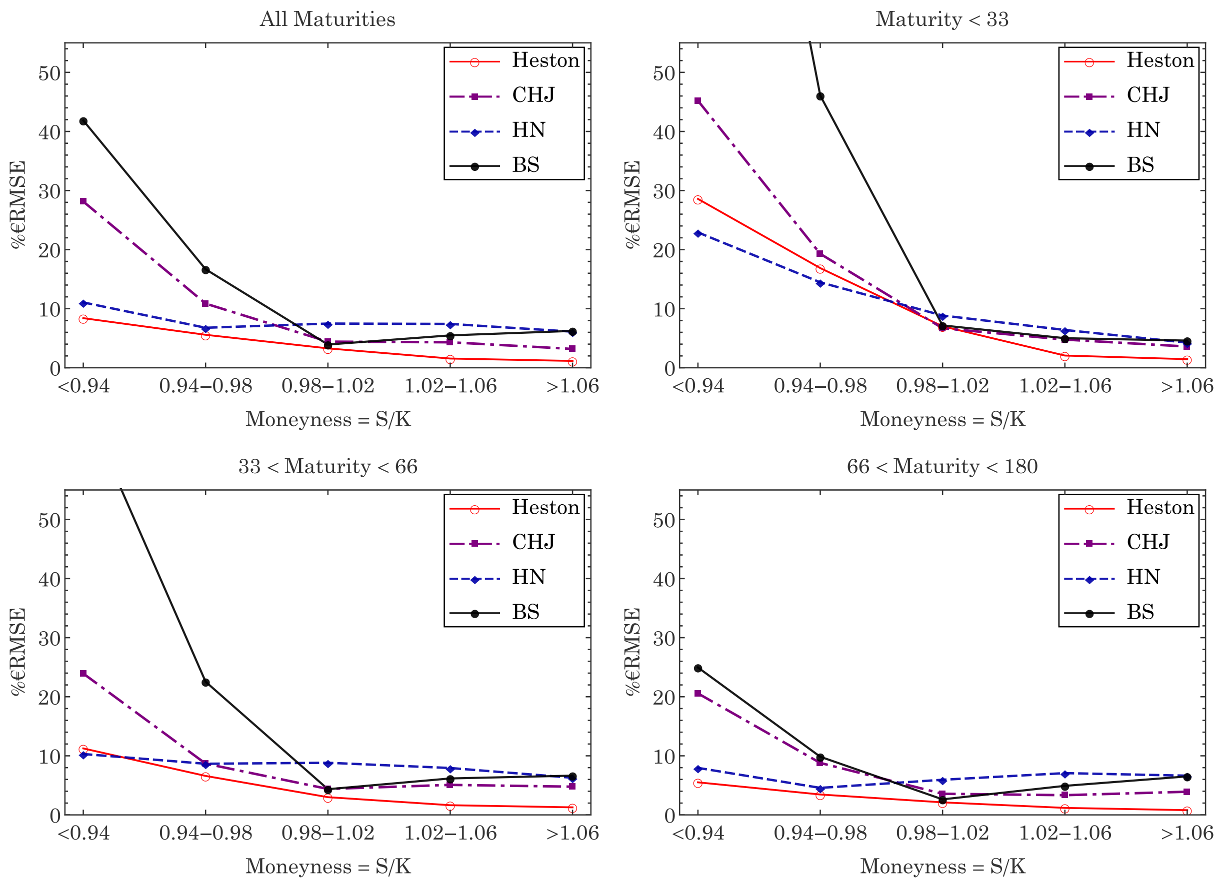

Figure 6 – €RMSE for the five moneyness categories produced by the models. Errors values are given in relation to the average price of each moneyness and TTM combination and displayed for all maturities together, as well as for each TTM separately.

Despite being calibrated with fairly equal weighting, if ever, more weight on short-term categories with higher moneyness, our models capture implied volatilities of longer-term options better. The HN model is, at least at some points, again even slightly better than the Heston model for OTM calls. However, what becomes clear in the last column and especially for contracts with less than 33 days to maturity is the inability of the model to capture the option skew (see also Figure 7). With its given parameters it produces an IVRMSE for all maturities of up to 6.52%, barely beating the BSM model at 7.34%. Moving to the CHJ model we can see in the results that just like for the €RMSE optimization it improves on the HN approach by, figuratively speaking, tilting the error curve. It is especially able to match the skew better, what results in a fairly satisfactory fit and an almost symmetrical percentage error curve, lowest for ATM contracts. Finally, the BSM model leads to almost symmetrical errors around the spot price as well. Its errors are mostly higher than in the other models and due to the constant volatility assumption the BSM model has especially trouble fitting the steeper slope of our first TTM group. With the flattening of the volatility smile the BSM fares somewhat better for longer maturities, even so, overall it is still worst.

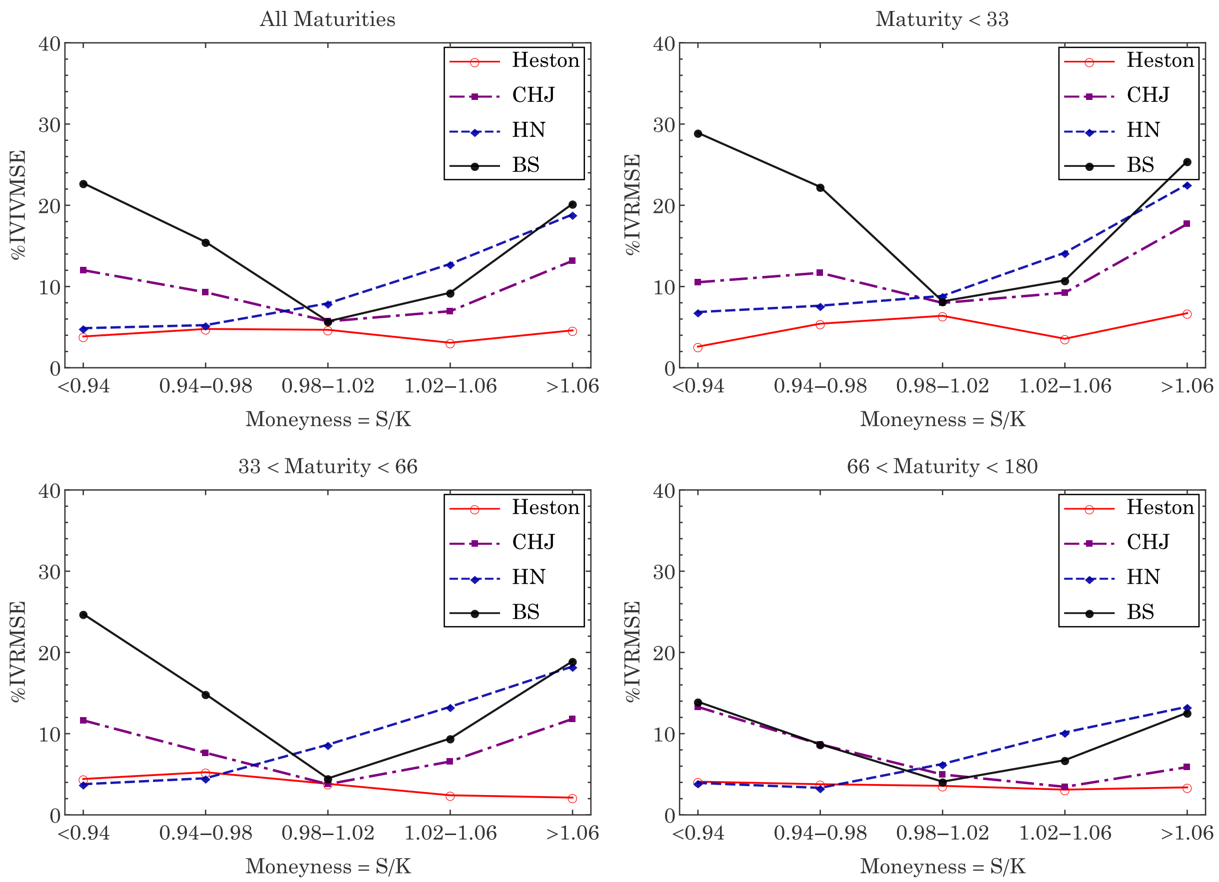

Figure 7 – IVRMSE for the five moneyness categories produced by the models. Errors values are given in relation to the average implied volatility of each moneyness and TTM combination and displayed for all maturities together, as well as for each TTM separately.

Aggregating over moneyness yields only limited additional value. For the €RMSE function errors are relatively flat and ranked just as total results are. The same hold for the IVRMSE objective function, however, the relative errors converge for longer maturities. Looking at Figures 6 and 7 again, we can see that a combination of the HN and CHJ models could seem like an ideal choice. The strength of the HN model for lower moneyness categories and for the CHJ for higher moneyness is evident across all maturities. Following a slightly different suggestion by Andries et al. (2015), who estimate the price of variance risk and show a decreasing premium for longer maturities, we estimate a separate ξ parameter for each maturity category. The calibrations indeed yield a lower premium for the longest TTM but the resulting improvements are fairly small in our case and therefore not reported. This indicates that we're already close to the optimum, given the limitations of our parameter estimates from returns. These alone do obviously not generate enough skewness to match the risk-neutral distribution, despite the mapping. On the other hand, for OTM calls and especially short-term contracts, which are those with lowest implied volatilities in our sample, we see a fairly satisfactory fit.

3.4 Out-of-Sample Results

While the flexibility to fit the observed data is a meaningful criterion for an option valuation model it is much more important to have a good out-of-sample performance. The ability to predict prices is a central aspect since the goodness-of-fit could be high due to overfitting, despite the model being misspecified. Evidence for GARCH models is provided by Christoffersen and Jacobs (2004b), who find that a rather simple model, allowing for the leverage effect and volatility clustering, trails more sophisticated models in-sample but outperforms these out-of-sample. Similarly, Dumas et al. (1998) demonstrate that a better in-sample performance due to increased flexibility does not necessarily translate into satisfactory predictive power. Capturing the most relevant stylized facts without fitting to temporary noise seems to lead to good performance, especially when no recalibration is conducted. This is also confirmed by Christoffersen et al. (2006). Since their GARCH model with inverse Gaussian innovations nests the HN model it obviously provides a better fit. However, in out-of-sample valuations it only performs better for OTM put options and even this is only the case when the model parameters are not held constant for more than about 10 weeks.

We keep all parameters from the calibration as they are but we allow the BSM model to update volatility and the Heston model to obtain a fresh initial variance for each cross-section. The estimates are presented in Table 8 and do not differ much between the two objective functions. Yet, note the stronger increase in weeks 7 and 8 for the IVRMSE estimated values of the BSM model. These indicate not only a general surge in market implied volatility but, compared to the €RMSE values of the same weeks, they suggest stronger volatility increases for shorter maturities. Surprisingly the same is not observable for the Heston model. Recall however that the Heston model has a different set of time-invariant parameters for each of the objective functions. Furthermore, for all six weeks of sample A, converting the Heston initial variances by taking the square root leads to almost identical values as the BSM implied volatilities. Moving to Sample B the Heston model displays a much stronger adjustment by more than doubling on the previous values to roughly 0.1052, which is equal to a standard deviation of 32.43%.

	Sample A						Sample B	
	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6	Week 7	Week 8
BSM – σ								
€RMSE	0.2288	0.2204	0.2195	0.2003	0.2050	0.2037	0.2694	0.2620
IVRMSE	0.2339	0.2198	0.2205	0.1983	0.2003	0.1995	0.2904	0.2951
Heston – v								
€RMSE	0.0677	0.0568	0.0569	0.0392	0.0435	0.0435	0.1052	0.0990
IVRMSE	0.0655	0.0543	0.0545	0.0365	0.0414	0.0415	0.1047	0.0989

Table 8 – BSM implied volatility and Heston initial variance updates for the out-of-sample periods.

The HN and CHJ models predict the out-of-sample weeks with only the conditional variance providing new information from daily updating through the return data. No further estimations or updates are necessary, not even for the independent variance premium parameter, which makes it a very convenient approach. According to Heston and Nandi (2000), alternatively treating the conditional variance as an additional parameter does not improve the model performance, aside from better in-sample fitting.

The errors for the 990 contracts of sample A are given in Table 9 and are split up by moneyness and maturity for each objective function. Bold values, which can be found in the last rows are aggregate errors for the entire sample. Beginning with the analysis of the €RMSE results we can see that the Heston model does exceptionally well for the medium and long maturity categories. Its aggregate error is just 2.15% of the average option price which can be found in Table 11 and is a slightly better value than it achieved in-sample. A much bigger improvement is noticeable for the GARCH models. With 4.29% for the CHJ model and 5.35% for the HN model these errors are still about twice as high as those of the Heston model, nonetheless, they are much better than their in-sample counterparts. Furthermore, they are much steadier across the different categories. Still worst in comparison is the BSM model, but it is seeing some improvements as well. The lower errors for all our models can be attributed to the relatively calm period and the less pronounced skew in data particularly helps the GARCH models.

Looking at the IVRMSE results we see a slightly different ranking of the models. Especially for the shortest maturity and for low moneyness the Heston model has trouble matching market data. These are also the categories which less emphasis is put on when recalibrating the model (see Table 11) and for some it is even outperformed by the BSM model. The CHJ model surprisingly outperforms all others, despite unchanged parameters. It can be seen in the last column of Table 9 that it produces rather low deviations from market data for the higher moneyness categories. What we can also notice is its good overall performance which stems from its continuity and steadiness. It barely ever provides the best fit to a data category, however, it also never fails as drastically as the other models do for some nodes. Plotting the relative errors, we can observe the same for relative values in Figure 8. With a total absolute total error of 2.27% the original HN model is just slightly worse. It clearly outperforms the BSM model and leads to results remarkably close to the Heston model. While it does well across all maturity categories it has, just like in-sample, some trouble fitting OTM puts and generally options far from the strike. Aside, it matches the shortest maturity almost consistently best. For the sake of completeness, we check the errors of the BSM model. Just like in all our other results, due to the underlying model structure, it displays a decent fit to ATM contracts for medium maturity. Lastly, unreported results show that equipping the CHJ model with a separate ξ parameter for each maturity leads to improvements over the HN model. Yet, the relevance of this method is questionable since it does not outperform the CHJ model in our out-of-sample analysis.

Out – of – Sample Valuation Errors Disaggregated by Moneyness and Maturity

Sample A		€RMSE				IVRMSE			
		< 33	33 – 66	66 – 180	All	< 33	33 – 66	66 – 180	All
< 0.94	Heston	4.29	6.38	8.95	6.73	4.89%	0.92%	0.60%	3.11%
	CHJ	6.84	20.30	42.91	27.46	1.96%	2.14%	2.47%	2.19%
	HN	2.76	10.92	17.11	11.59	3.73%	0.93%	0.81%	2.43%
	BSM	20.40	39.28	56.89	50.94	4.10%	3.30%	2.59%	3.43%
0.94 – 0.98	Heston	12.66	11.70	9.88	11.46	1.49%	1.43%	0.65%	1.23%
	CHJ	12.31	20.01	32.96	23.70	2.32%	1.60%	1.82%	1.99%
	HN	6.88	18.70	17.33	14.53	1.24%	1.23%	0.71%	1.07%
	BSM	27.04	28.91	36.56	31.37	3.94%	1.81%	1.47%	2.77%
0.98 – 1.02	Heston	24.59	10.57	8.80	17.27	2.77%	1.12%	0.51%	1.86%
	CHJ	15.99	20.46	22.47	19.59	2.42%	1.19%	1.20%	1.79%
	HN	15.15	32.03	32.26	26.65	1.73%	1.94%	1.26%	1.63%
	BSM	22.94	17.92	13.65	18.86	2.78%	0.93%	0.64%	1.87%
1.02 – 1.06	Heston	18.31	6.77	7.59	13.13	3.30%	0.57%	0.49%	2.19%
	CHJ	11.55	26.12	24.68	20.64	1.93%	1.39%	0.84%	1.52%
	HN	18.15	42.30	50.62	37.62	2.01%	2.81%	2.09%	2.26%
	BSM	19.70	39.52	39.15	32.47	2.31%	2.68%	1.82%	2.27%
> 1.06	Heston	8.79	5.44	8.83	8.10	2.72%	0.60%	0.88%	1.85%
	CHJ	11.74	30.88	35.40	26.76	2.09%	2.16%	1.28%	1.88%
	HN	17.78	44.58	61.25	43.45	3.26%	3.80%	2.97%	3.31%
	BSM	24.96	54.85	66.43	49.92	5.11%	4.84%	3.47%	4.55%
All	Heston	15.44	8.54	8.93	**11.94**	3.11%	0.98%	0.65%	**2.08%**
	CHJ	12.17	24.12	32.43	**23.85**	2.16%	1.75%	1.62%	**1.89%**
	HN	13.78	32.72	39.46	**29.70**	2.53%	2.43%	1.77%	**2.27%**
	BSM	23.44	38.57	45.92	**36.36**	3.83%	3.07%	2.21%	**3.16%**

Table 9 – Detailed and aggregate out-of-sample RMSE values for both objective functions for the six weeks of sample A.

As noted, the average values of Sample A, which are presented in Table 11, point to an even less turbulent time than our in-sample data. Average prices are slightly lower, especially for the longest maturity and the implied volatility surface flattens notably. Using these averages again we calculate the percentage deviations of each model for both objective functions and plot them for the different moneyness categories. The relative errors are given in Figure 8. To facilitate comparison across the maturities and to distinguish the model performances better, we apply the same scale to our results and cut off some outliers produced by the BSM model. The characteristics are similar to the in-sample data. To match longer maturities and high moneyness categories all sequentially or purely option estimated models sacrifice some performance for lower priced contracts. The tilt of the CHJ error curve compared to the HN model is also noticeable again. Furthermore, the percentage error curves of the CHJ model resemble those in Figure 7, allowing us to infer

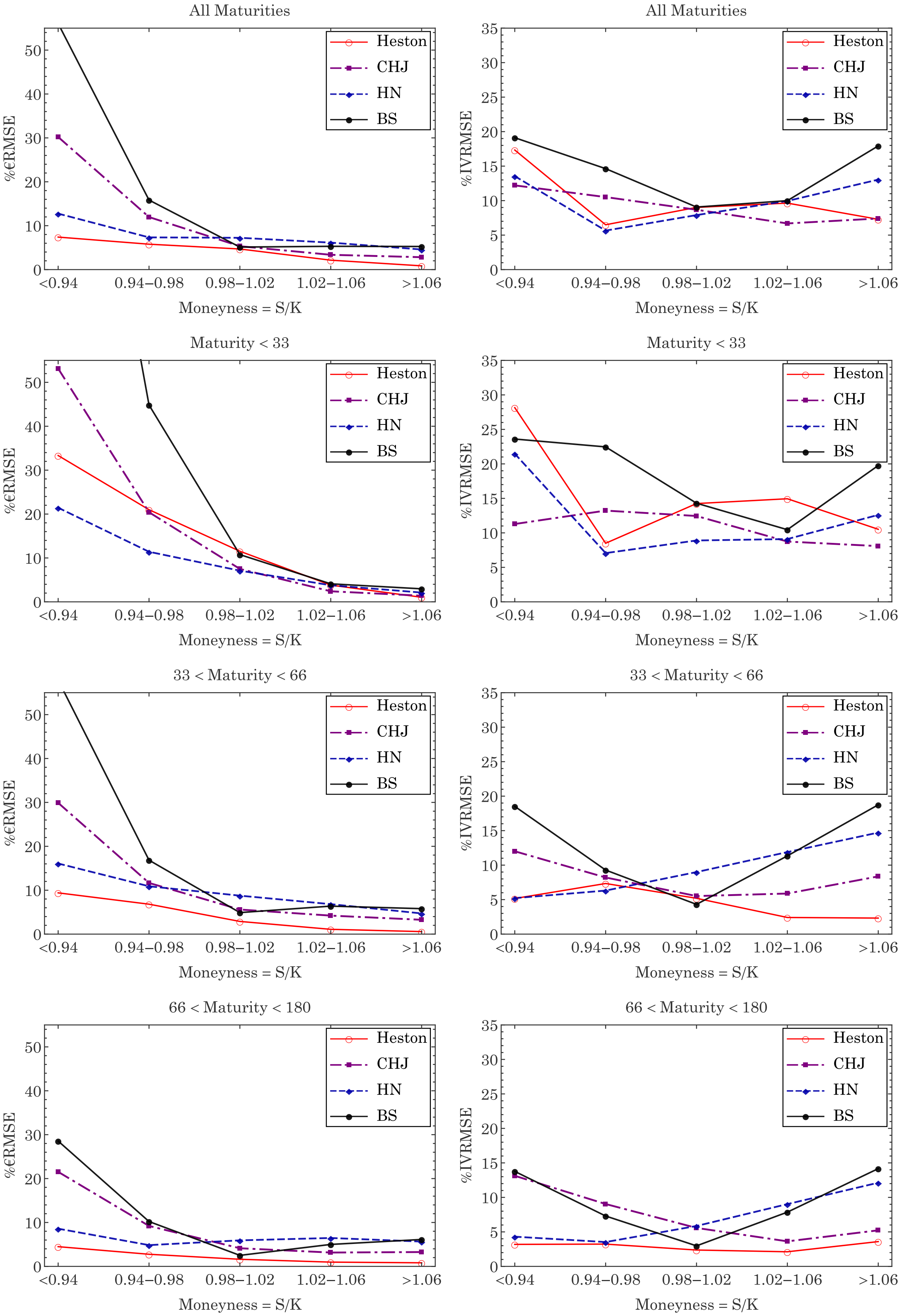

Figure 8 – Percentage out-of-sample errors for sample A using two different objective functions. These are obtained by dividing the model errors through the average prices and implied volatility values of each category.

an absence of overfitting. We also get a graphical representation of the HN model doing really well for short maturities. This translates to an overall good explanation of sample A data, outperforming the Heston model for the IVRMSE objective function in almost all cases. While calculating relative results doesn't change the ranking of our models they still yield plenty of additional information. Ultimately, for investors the return, i.e., relative performance, is of importance. For instance, a €RMSE for the shortest TTM of slightly above 20€, produced by the BSM model, has totally different implications depending on the moneyness of an option. For OTM calls the corresponding average market price is 12.87€, which results in an error of about 155%. At the same time the relative error for a deep ITM call with average price of close to 200€ is much lower as it is only around 10%.

The more interesting results, however, are, just as seen for the absolute error values in Table 9, those of the IVRMSE minimized objective function. Applying the Heston model leads to an aggregate error more than twice as high as in-sample. In the representation of percentage errors, it becomes even clearer that the Heston model has trouble matching the shortest maturity, producing errors of up the 30%. Apparently the tight in-sample fit is achieved mostly by overfitting the data. In contrast to that the GARCH models even improve on their fitting performance. It seems they indeed capture the most relevant characteristics without attempting to match temporary fluctuations which do not persist. Calculating the mispricings from market implied volatility for all contracts gives percentage IVRMSE values of 9.77% for the Heston model, 8.91% and 10.71% for the two GARCH models and 14.91% for the BSM model.

This slightly better than in-sample performance for the BSM model points to a flatter volatility surface and also serves as a reasonable explanation for the overall weaker performance of the Heston model. The state variable ν_t can only affect the level but not the slope of the volatility skew at a given maturity. Even with an optimized value each week it cannot reproduce the flatter curve. Depending on the conditional variance, the GARCH models are able to generate differently sloped implied volatility curves for the same parameter set. They even affect the slope of certain moneyness categories stronger than others. For our estimates, but also for the jointly estimated parameters of Heston and Nandi (2000), an increase in conditional variance has the biggest impact on short maturities and OTM calls. With an arbitrary TTM of 30 days the difference between low and high conditional variance peaks at a moneyness of 0.9 to 1.00. Including a positive ξ emphasizes this effect further but does not change the location of the peak.

Proceeding to the shorter and more volatile remainder of our forecasting period, we evaluate the models on two weeks of data leading up to the UK referendum. The sub-sample comprising 289 contracts displays entirely different characteristics than Sample A (see Table 11). Generally, average option prices are higher and implied volatility almost doubled in the most extreme instance. While some of the option price increases can certainly be attributed to TTM changes, i.e., the June contracts drop out due to our expiry threshold of at least 7 days, most is caused by the substantial rise in implied volatility. As a consequence, the previously almost flat term structure steepens notably and the skew along the moneyness axis is much more pronounced as well. We can also see that most of the increase stems from investors seeking downside protection against the upcoming event in OTM puts with relatively short maturities. The average implied volatilities of the longest maturity in our sample are barely affected. With increases between just 2% and 3% market participants on average seem to expect the turbulences to settle within a few weeks.

The aggregate percentage errors under these changed circumstances for both objective functions are presented in Figure 9. In direct comparison the Heston model still does relatively well, despite percentage RMSEs multiple times higher than in-sample. Compared to Sample A, however, all models perform much worse. What's most striking are the significantly higher errors for the GARCH models, which are at many instances even outperformed by the BSM model. Inadequately capturing prices and not exhibiting sufficient skew across all dimensions, this failure can largely be explained by looking at Figure 10. For the relatively calm Sample A, which ends at the first vertical black line, we observe an annualized conditional volatility which behaves fairly similar to the market implied volatility shown in the last row of the same Figure. However, towards the end of Sample B, indicated by the second vertical black line, conditional volatility declines in accordance with positive returns. At the same time implied volatility, given each Wednesday, moves in the opposite direction, which causes the correlation to break. In anticipation of the upcoming event, represented by the red line, market implied volatility increases to its highest level in our out-of-sample period, consequently driving option prices up. Since this information is not yet reflected in the dynamics of the underlying index returns, especially the HN model has no possibility to adjust to the generally higher level of volatility. Ultimately, the model then predicts option prices based on outdated information which leads to consistently lower than observed results. Furthermore, as we've

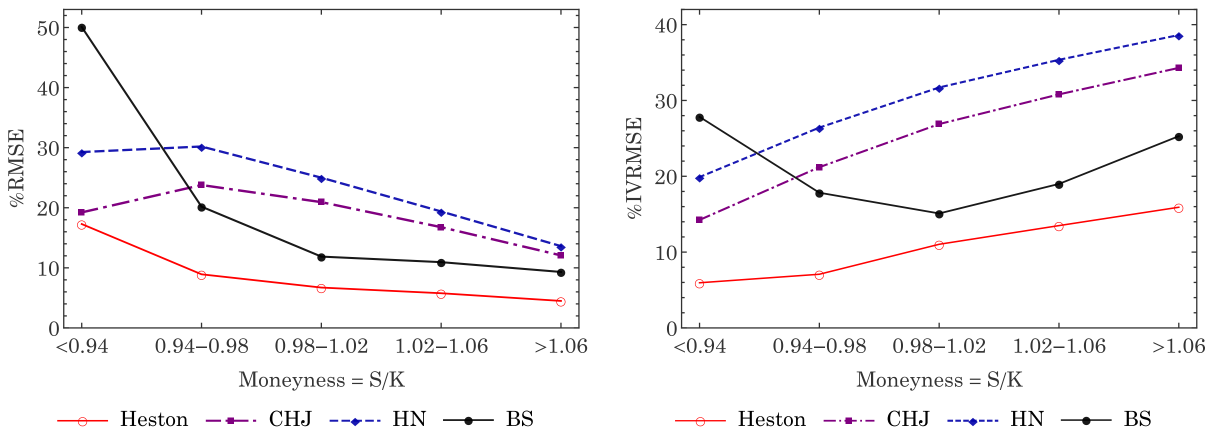

Figure 9 – Sample B percentage RMSE values, given in relation to average market data for different moneyness categories.

seen in Figure 5, the skewness parameter enables the conditional variance to react stronger to movements of the underlying. When looking at option based or joint calibrations we find that our results allow for comparably little skewness. In contrast, these calibrations often yield much higher parameter estimates for γ^* than purely return based approaches (Heston and Nandi, 2000; Christoffersen et al., 2013).

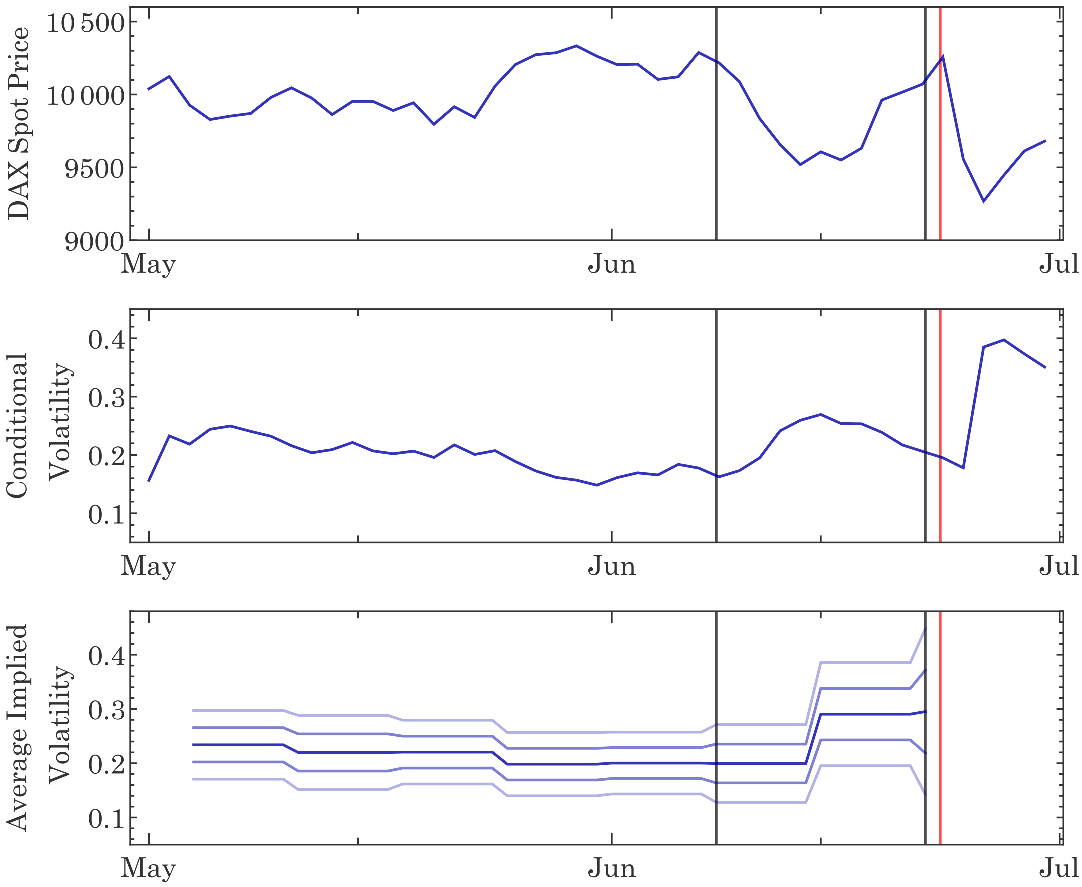

Figure 10 – DAX spot prices, annualized conditional volatility and average annualized implied volatility with bands for one and two standard deviations. The first vertical black line indicates the end of sample A and the beginning of sample B, which lasts up to the seconds black line. The red line highlights the day of the UK referendum.

While a re-estimation of the latent variable during such a temporary volatility surge leads to somewhat acceptable results in the Heston model, the GARCH models would require an entire recalibration of all parameters under such circumstances. Unless the model is frequently updated, this could also not be overcome by incorporating option data in the estimation. Unlike the instantaneous variance, the design of the GARCH models leads to a delayed reaction in the conditional variance. As we can observe in Figure 10 the DAX index drops by several hundred points on the day following the referendum. At the same time, however, the conditional volatility still declines based on the positive return of the previous day and doesn't increase until the next step in time. For the GARCH models to capture this situation we would first have to fit the parameters in the second half of sample B, linking a smoothly increasing underlying to higher implied volatilities. While this would provide us with fairly decent estimates for one week, once the index drops and conditional variance, or annualized conditional volatility as displayed in Figure 10, increases the model parameters would have to be recalibrated once again to fit the new situation.

Yet, the tendency is that GARCH models are used with long time-series to establish a reliable connection between conditional variance and the data, irrespective if options, returns or both. This procedure leads to somewhat averaged parameters, which might not capture all market anomalies, but fit the general structure very well. In our analysis this insight is reinforced by the return based estimation. The in-sample fit of our GARCH models is mediocre, however, they do not suffer from overfitting and their performance even improves when forecasting Sample A. Regarding the delayed reaction in conditional variance it might be useful to reduce the length of time steps. However, no matter how small the increment, the delay is still present. Since the GARCH models need at least one updated return information they can at soonest react two time steps after an event.

4 Conclusion

In this work we conduct an empirical analysis to evaluate the model of Heston and Nandi (2000). We begin by highlighting the desired properties a suitable option pricing model should display and by investigating the HN GARCH dynamics of the variance process. For comparison and to gain an understanding of the performance of the HN model, we also describe and provide results for other widely used and well-known models. Namely the BSM model and the Heston model, which serve us as benchmarks since their behavior is comparably well understood. Lastly we introduce a more recent development in GARCH option pricing literature suggested by CHJ (2013). Their proposal nests the HN model and therefore results in a better in-sample fit. However, since they do not conduct an out-of-sample analysis, comparing their extension to the original HN model, it is included as well. We obtain GARCH model parameters through maximum likelihood estimation on a series of returns, using an alternative and computationally more efficient procedure. For the model comparison we additionally include option data and distinguish between two objective functions. Furthermore, we emphasize and show the importance of the objective function by reporting extensive results on multiple cross-sections in-sample, as well as for two differently characterized out-of-sample periods.

Our empirical results are instructive in a number of ways. As expected, the Heston model fares by far best when it comes to fitting. However, out-of-sample and especially when using the parameter sets obtained by optimizing the IVRMSE, the GARCH models get surprisingly close. This is remarkable because the HN model is fully and even the CHJ model is largely based on return estimated parameters. Furthermore, they are not updated or optimized in any way, aside from the information which stems from the underlying index returns. In general, our results indicate that the GARCH models do not suffer from overfitting and are able to capture market dynamics associated with stylized facts such as volatility clustering or the leverage effect. These findings hold true, irrespective of specifying an independent variance risk premium or not. Nonetheless, in our case the CHJ (2013) model at almost all instances performs better than the HN model. An attempt to further improve performance by giving flexibility to said premium for different maturities does not lead to notably better results. Yet, including a more flexible risk premium might prove useful when calibrating the model to a more volatile in-sample period with a steeper term structure of implied volatility.

Altogether, it is satisfying to see the results in line with the theoretical expectations from Chapter 2. Particularly the properties, physical and risk-neutral, point to consistent and plausible findings. Most of the performance of the GARCH models, but especially the HN model, does in fact stem from the ability to simultaneously consider the path dependence and negative correlation of volatility with returns. Furthermore, our analysis shows that a U-shaped pricing kernel not only fits the data better, it also results in improved out-of-sample performance in both of our forecasting periods.

We can also conclude that the GARCH models exhibit an absolutely satisfactory performance and do not require much attention once estimated. Given a chance to capture general market dynamics this works very well for uneventful periods. While in our case the GARCH models did not necessarily fail for shorter maturities, calibrating them to predict option prices with longer TTMs on rather steady underlyings might be a sensible scope of application. Else, in times of frequent fiscal and especially monetary interventions at prespecified dates, the performance of the GARCH models could suffer.

This weakness becomes apparent in one of our main findings, which is the inability of the GARCH models to adapt to implied volatility changes in anticipation of scheduled events. The increased demand for insurance and possibly additional speculation affects option prices and is not necessarily reflected in return data. However, even when calibrating the models jointly by including option data, two weeks of slightly raised implied volatility levels would not have a big impact in a sample which is several months or years long. Since empirical studies often employ large amounts of data, they might easily miss our findings in Sample B. A few cross-sections of bad performance quickly smoothen out when dozens or hundreds of observations are made. Furthermore, due to the structure of the model the delayed reaction is always present, even with shorter time steps.

A conceivable way to generally correct the lack of conditional variance, as observed in our sample B period, could be the attempt to incorporate non-normal innovations. Also, a GJR-GARCH or EGARCH like formulation might prove fruitful when it comes to introducing the necessary skewness, even if at the cost of an almost closed-form solution. It would be interesting to see how the GARCH models and especially the one with augmented pricing kernel perform out-of-sample when calibrated to a joint set of option and return data. Moreover, comparing the discrete-time models of this work to a continuous model, which does not treat initial variance as parametric, should provide interesting results.

A The Black-Scholes-Merton Differential Equation

The price of an option is a function of its underlyings price and the time. In the BSM scenario, the underlying evolves according to equation (2.2). Our goal is to construct a risk-free portfolio Ψ, consisting of Δ times the underlying S_t and the option $V(S_t, t)$, using delta hedging. Taking a long position in S_t and shorting the option $V(S_t, t)$ yields

$$\Psi_t = \Delta S_t - V(S_t, t), \tag{A.1}$$

with change in value over an infinitesimal time step according to

$$d\Psi_t = \Delta dS_t - dV(S_t, t). \tag{A.2}$$

By applying Itô's Lemma to our process and simply writing V for $V(S_t, t)$, it follows that

$$d\Psi_t = \Delta dS_t - \left(\frac{\partial V}{\partial t}dt + \frac{\partial V}{\partial S_t}dS_t + \frac{1}{2}\frac{\partial^2 V}{\partial S_t^2}dS_t^2\right) \tag{A.3}$$

$$= \Delta dS_t - \left(\frac{\partial V}{\partial t}dt + \frac{\partial V}{\partial S_t}dS_t + \frac{1}{2}\sigma^2 S_t^2\frac{\partial^2 V}{\partial S_t^2}dt\right) \tag{A.4}$$

$$= \left(\Delta - \frac{\partial V}{\partial S_t}\right)dS_t - \left(\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S_t^2\frac{\partial^2 V}{\partial S_t^2}\right)dt. \tag{A.5}$$

Setting $\Delta - \frac{\partial V}{\partial S_t} = 0$ by choosing $\Delta = \frac{\partial V}{\partial S_t}$ and continuously rebalancing this relationship equals a dynamic delta hedge. Since the value of the portfolio is now independent of the Wiener process W_t, it changes deterministically over the next time step. And as the no-arbitrage condition is imposed, this change must be equal to the risk-free rate of return. Therefore,

$$d\Psi_t = r\Psi_t dt, \tag{A.6}$$

with r representing the risk-free rate. Substituting the preceding expression and (A.1) yields

$$\left(\frac{\partial V}{\partial S_t} - \frac{\partial V}{\partial S_t}\right)dS_t - \left(\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S_t^2\frac{\partial^2 V}{\partial S_t^2}\right)dt = r\left(\frac{\partial V}{\partial S_t} - V(S_t, t)\right)dt \tag{A.7}$$

and simplifying

$$rV = \frac{\partial V}{\partial t} + rS_t\frac{\partial V}{\partial S_t} + \frac{1}{2}\sigma^2 S_t^2\frac{\partial^2 V}{\partial S_t^2}. \tag{A.8}$$

This is the well-known BSM partial differential equation, which does no longer include any market price of risk. All remaining variables, S_t, σ and r are independent of investors' risk preferences.

B Deriving the Heston Risk-Neutral Dynamics

In order to obtain the risk-neutral versions (2.12) and (2.13) from equations (2.9) and (2.10) we use Itô's Lemma again. Rewriting equation (2.9) to

$$\frac{dS_t}{S_t} = \upsilon dt + \sqrt{\nu_t} dW_t^{(1)} \tag{B.1}$$

and applying Itô's Lemma to $x_t = g\left(S_t, t\right) = \log\left(S_t\right)$, since $\frac{dS_t}{S_t}$ is the rate of return over the next infinitesimal short time interval, yields

$$\frac{\partial g}{\partial t}\left(S_t, t\right) = 0, \qquad \frac{\partial g}{\partial S_t}\left(S_t, t\right) = \frac{1}{S_t}, \qquad \frac{\partial^2 g}{\partial S_t^2} = -\frac{1}{S_t^2},$$

$$d\log\left(S_t\right) = \left(0 + \frac{1}{S_t}\upsilon S_t + \frac{1}{2}\left(-\frac{1}{S_t^2}\right)\nu_t S_t^2\right) dt + \frac{1}{S_t}\sqrt{\nu_t} S_t dW_t^{(1)} \tag{B.2}$$

$$= \upsilon dt + \sqrt{\nu_t} dW_t^{(1)} - \frac{1}{2}\nu_t dt \tag{B.3}$$

$$dx_t = \left(\upsilon - \frac{1}{2}\nu_t\right) dt + \sqrt{\nu_t} dW_t^{(1)}. \tag{B.4}$$

Writing (B.4) as the risk-neutral process

$$dx_t = \left(r - \frac{1}{2}\nu_t\right) dt + \left(\upsilon - r\right) dt + \sqrt{\nu_t} dW_t^{(1)} \tag{B.5}$$

$$= \left(r - \frac{1}{2}\nu_t\right) dt + \sqrt{\nu_t} dW_t^{*(1)}, \tag{B.6}$$

where

$$W_t^{*(1)} = W_t^{(1)} + \frac{\upsilon - r}{\sqrt{\nu_t}} t, \tag{B.7}$$

is equal to equation (2.12).

Following Rouah and Vainberg (2007), the risk-neutral dynamics of (2.10) are given by

$$d\nu_t = \kappa\left(\theta - \nu_t\right) dt - \lambda dt + \varrho\sqrt{\nu_t} dW_t^{*(2)}, \tag{B.8}$$

where

$$W_t^{*(2)} = W_t^{(2)} + \frac{\lambda}{\varrho} t. \tag{B.9}$$

Using the price of volatility risk given in Heston (1993), $\lambda(S, \nu, t) = \lambda\nu_t$, yields

$$d\nu_t = \kappa(\theta - \nu_t)\,dt - \lambda\nu_t dt + \varrho\sqrt{\nu_t} dW_t^{*(2)}. \tag{B.10}$$

For $\kappa^* = \kappa + \lambda$ and $\theta^* = \frac{\kappa\theta}{\kappa+\lambda}$ this is equal to equation (2.13) since it reduces to

$$d\nu_t = \kappa^*(\theta^* - \nu_t)\,dt + \varrho\sqrt{\nu_t} dW_t^{*(2)}. \tag{B.11}$$

C Discontinuities in the Heston (1993) Model

The parameters used to highlight the occurring discontinuities in the Heston (1993) CF are taken from the in-sample calibration. Of these, the time-invariant parameters are $\varrho = 1.0214$, $\rho = -0.6948$, $\kappa = 3.7499$ and $\Theta = 0.0660$, with the initial variance set equal to the first week's estimation $\nu_1 = 0.0683$ ($\approx 26.14\%$ annualized volatility) and $\tau = \frac{T-t}{252}$ set to the maturity of the December contract at $\frac{176}{252} = \frac{44}{63} \approx 0.7$ years.

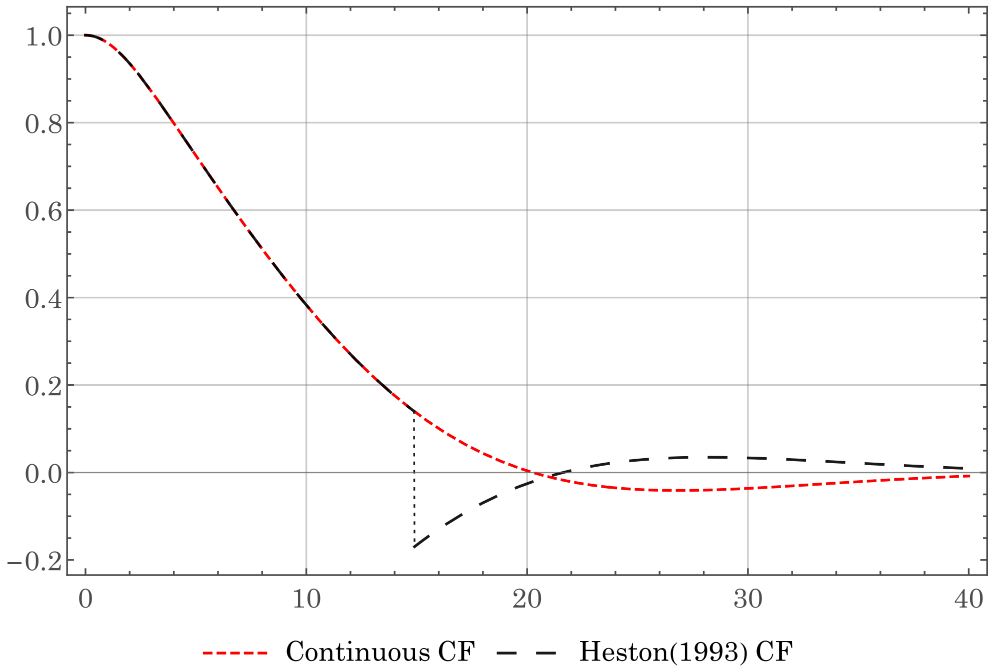

Figure 11 – Continuous and discontinuous formulation of the Heston (1993) characteristic function.

As can be seen in Figure 11, the discontinuity occurs at a value of about $\phi = 14.89$. With the continuous CF, the option is assigned a price of 7.5869. The mispricing in the original formulation amounts to 1.9469, since the erroneous price is 9.5338. This equals a deviation from the considered correct price of about 25.66%.

D Central Moments and Dynamics of the Heston & Nandi (2000) Model

In the HNGARCH model the first conditional moment of returns for $z_t \overset{iid}{\sim} \mathcal{N}(0,1)$ is given by

$$\mathbb{E}[R_t] = \mathbb{E}\left[r + \left(\mu - \frac{1}{2}\right) h_t + \sqrt{h_t} z_t\right] \tag{D.1}$$

$$= r + \left(\mu - \frac{1}{2}\right) h_t. \tag{D.2}$$

The second conditional moment of returns is

$$Var(R_t) = Var\left(r + \left(\mu - \frac{1}{2}\right) h_t + \sqrt{h_t} z_t\right) \tag{D.3}$$

$$= Var\left(\sqrt{h_t} z_t\right) \tag{D.4}$$

$$= h_t. \tag{D.5}$$

The unconditional variance can be obtained by rewriting equation (2.28) as follows,

$$h_{t+1} = \omega + \alpha + \left(\beta + \alpha\gamma^2\right) h_t + \alpha\left(\left(z_t - \gamma\sqrt{h_t}\right)^2 - 1 - \gamma^2 h_t\right). \tag{D.6}$$

Assuming $\left(\left(z_t - \gamma\sqrt{h_t}\right)^2 - 1 - \gamma^2 h_t\right)$ to be a zero-mean innovation, as in Christoffersen et al. (2014) and given that the process remains stationary, which is the case for $\beta + \alpha\gamma^2 < 1$, the unconditional variance can quickly be derived. This gives

$$\mathbb{E}[h_t] = \frac{\omega + \alpha}{1 - \beta - \alpha\gamma^2}. \tag{D.7}$$

The variance of the conditional variance is

$$Var(h_{t+1}) = Var\left(\omega + \beta h_t + \alpha\left(z_t - \gamma\sqrt{h_t}\right)^2\right) \tag{D.8}$$

$$= Var\left(\alpha\left(z_t - \gamma\sqrt{h_t}\right)^2\right) \tag{D.9}$$

$$= Var\left(\alpha z_t^2 - 2\alpha z_t \gamma\sqrt{h_t}\right). \tag{D.10}$$

Since z_t^2 is chi-squared distributed with $k = 1$ degree of freedom it holds that $Var\left(z_t^2\right) = 2k = 2$, which yields

$$Var(h_{t+1}) = 2\alpha^2 + 4\alpha^2\gamma^2 h_t \tag{D.11}$$

$$= 2\alpha^2\left(1 + 2\gamma^2 h_t\right). \tag{D.12}$$

E Derivation of the Log-Likelihood Function

From equations (2.25) and (2.26), or alternatively from (D.2) and (D.5) we know that the returns are normally distributed with $R_t \sim \mathcal{N}\left(r + \left(\mu - \frac{1}{2}\right) h_t, h_t\right)$. By setting the first observation h_0 to either the sample variance or, as in our case, to the variance of the previous 15 trading days, we can obtain the parameters by maximizing an appropriate likelihood function. Due to the independent distribution of innovations this is given by

$$\mathcal{L}(\alpha, \beta, \gamma, \omega, \mu) = \prod_{t=1}^{T} \frac{1}{\sqrt{2\pi h_t}} \exp\left(-\frac{\left(R_t - \left(r + \left(\mu - \frac{1}{2}\right) h_t\right)\right)^2}{2h_t}\right). \tag{E.1}$$

Since taking the logarithm of the likelihood function does not affect the location of the maximum, we will get the same parameter estimates by writing

$$\log \mathcal{L}(\alpha, \beta, \gamma, \omega, \mu) = \log\left(\prod_{t=1}^{T} \frac{1}{\sqrt{2\pi h_t}} \exp\left(-\frac{\left(R_t - \left(r + \left(\mu - \frac{1}{2}\right) h_t\right)\right)^2}{2h_t}\right)\right), \tag{E.2}$$

what simplifies calculation, since it reduces to

$$\log \mathcal{L}(\alpha, \beta, \gamma, \omega, \mu) = \sum_{t=1}^{T} \left(-\log\left(\sqrt{2\pi h_t}\right) - \frac{\left(R_t - \left(r + \left(\mu - \frac{1}{2}\right) h_t\right)\right)^2}{2h_t}\right) \tag{E.3}$$

$$= \sum_{t=1}^{T} \left(-\frac{1}{2}\log(2\pi) - \frac{1}{2}\log(h_t) - \frac{1}{2}\left(\underbrace{\frac{R_t - \left(r + \left(\mu - \frac{1}{2}\right) h_t\right)}{\sqrt{h_t}}}_{z_t}\right)^2\right) \tag{E.4}$$

$$= -\frac{T}{2}\log(2\pi) - \frac{1}{2}\sum_{t=1}^{T}\left(\log(h_t) + z_t^2\right). \tag{E.5}$$

Dropping the constant term, $-\frac{T}{2}\log(2\pi)$, yields the log-likelihood function, which is proportional to

$$\log \mathcal{L}(\alpha, \beta, \gamma, \omega, \mu) \propto -\frac{1}{2}\sum_{t=1}^{T}\left(\log(h_t) + z_t^2\right). \tag{E.6}$$

F Average Option Prices and Implied Volatilities

In – Sample

	Time to Maturity (days)	< 0.94	0.94 – 0.98	0.98 – 1.02	1.02 – 1.06	> 1.06	All
		Moneyness					
Average Price	< 33	16.06	69.23	237.33	505.88	847.77	343.11
	33 – 66	63.27	168.56	367.60	629.48	956.99	448.33
	66 – 180	251.92	413.83	628.98	884.19	1150.80	663.55
	All	112.45	223.50	419.81	668.11	986.91	**488.08**
Average Implied Volatility	< 33	16.91%	18.06%	20.79%	24.35%	28.87%	21.91%
	33 – 66	17.19%	18.84%	21.17%	23.59%	26.32%	21.53%
	66 – 180	18.81%	20.13%	21.52%	23.15%	24.49%	21.61%
	All	17.65%	19.05%	21.18%	23.70%	26.50%	**21.67%**

Table 10 – In-sample average prices and average implied volatilities for each moneyness and TTM combination.

Out – of – Sample

Sample A

	Time to Maturity (days)	< 0.94	0.94 – 0.98	0.98 – 1.02	1.02 – 1.06	> 1.06	All
		Moneyness					
Average Price	< 33	12.87	60.30	213.25	479.74	840.50	336.42
	33 – 66	67.88	171.60	365.55	619.72	945.69	445.20
	66 – 180	199.26	357.84	544.09	782.80	1083.77	593.47
	All	90.89	198.03	367.93	612.22	948.51	**452.55**
Average Implied Volatility	< 33	17.37%	17.54%	19.44%	22.08%	25.86%	20.61%
	33 – 66	17.82%	19.49%	21.57%	23.63%	25.82%	21.76%
	66 – 180	18.71%	20.04%	21.34%	22.85%	24.45%	21.48%
	All	17.94%	18.93%	20.63%	22.72%	25.38%	**21.19%**

Out – of – Sample

Sample B

	Time to Maturity (days)	< 0.94	0.94 – 0.98	0.98 – 1.02	1.02 – 1.06	> 1.06	All
		Moneyness					
Average Price	< 33	47.89	158.70	370.52	620.94	933.75	440.47
	33 – 66	136.92	278.96	487.54	738.66	1031.99	557.66
	66 – 180	291.57	441.33	674.20	886.02	1188.00	698.66
	All	145.67	277.36	503.26	738.14	1033.56	**555.39**
Average Implied Volatility	< 33	25.60%	29.95%	35.34%	39.82%	44.52%	35.34%
	33 – 66	22.60%	24.78%	27.72%	30.38%	33.20%	27.99%
	66 – 180	21.50%	22.78%	24.67%	26.06%	27.78%	24.59%
	All	23.17%	25.85%	29.01%	32.01%	35.31%	**29.28%**

Table 11 – Average out-of-sample prices, as well as average implied volatilities for each moneyness and TTM combination, given for the two subsamples.

Bibliography

Aït-Sahalia, Y., Kimmel, R., 2007. Maximum Likelihood Estimation of Stochastic Volatility Models. Journal of Financial Economics 83, 413–452.

Aït-Sahalia, Y., Lo, A. W., 1998. Nonparametric Estimation of State–Price Densities Implicit in Financial Asset Prices. The Journal of Finance 53, 499–547.

Albrecher, H., Mayer, P., Schoutens, W., Tistaert, J., 2006. The Little Heston Trap.

Amin, K. I., Ng, V. K., 1993. Option Valuation with Systematic Stochastic Volatility. The Journal of Finance 48, 881.

Andersen, L. B. G., 2007. Efficient Simulation of the Heston Stochastic Volatility Model.

Andries, M., Eisenbach, T. M., Schmalz, M. C., Wang, Y., 2015. The Term Structure of the Price of Variance Risk. FRB of New York, Working Paper No. FEDNSR736.

Attari, M., 2004. Option Pricing Using Fourier Transforms: A Numerically Efficient Simplification.

Bachelier, L., 1900. Théorie de la Spéculation. Gauthier-Villars.

Bakshi, G., Cao, C., Chen, Z., 1997. Empirical Performance of Alternative Option Pricing Models. The Journal of Finance 52, 2003–2049.

Bakshi, G., Madan, D., Panayotov, G., 2010. Returns of Claims on the Upside and the Viability of U-Shaped Pricing Kernels. Journal of Financial Economics 97, 130–154.

Barone-Adesi, G., Engle, R. F., Mancini, L., 2008. A GARCH Option Pricing Model with Filtered Historical Simulation. The Review of Financial Studies 21, 1223–1258.

Bates, D. S., 1996. 20 Testing Option Pricing Models. In: Maddala, G. S., Rao, C. R. (eds.), *Statistical Methods in Finance*, Elsevier, Amsterdam and New York, vol. 14 of *Handbook of Statistics*, pp. 567–611.

Bates, D. S., 2000. Post-'87 Crash Fears in the S&P 500 Futures Option Market. Journal of Econometrics 94, 181–238.

Bates, D. S., 2003. Empirical Option Pricing: A Retrospection. Journal of Econometrics 116, 387–404.

Black, F., 1976. Studies of Stock Price Volatility Changes. In: *Proceedings of the 1976 Meetings of the Business and Economics Statistics Section*, pp. 177–181.

Black, F., Scholes, M., 1973. The Pricing of Options and Corporate Liabilities. The Journal of Political Economy pp. 637–654.

Bollerslev, T., 1986. Generalized Autoregressive Conditional Heteroskedasticity. Journal of Econometrics 31, 307–327.

Bollerslev, T., Engle, R. F., 1993. Common Persistence in Conditional Variances. Econometrica: Journal of the Econometric Society pp. 167–186.

Brennan, M. J., 1979. The Pricing of Contingent Claims in Discrete Time Models. The Journal of Finance 34, 53–68.

Byun, S. J., 2011. A Study on Heston-Nandi GARCH Option Pricing Model. KAIST Business School, Korea.

Câmara, A., 2003. A Generalization of the Brennan-Rubinstein Approach for the Pricing of Derivatives. The Journal of Finance 58, 805–819.

Chernov, M., Ghysels, E., 2000. A Study Towards a Unified Approach to the Joint Estimation of Objective and Risk-Neutral Measures for the Purpose of Options Valuation. Journal of Financial Economics 56, 407–458.

Christie, A. A., 1982. The Stochastic Behavior of Common Stock Variances: Value, Leverage and Interest Rate Effects. Journal of Financial Economics 10, 407–432.

Christoffersen, P., Elkamhi, R., Feunou, B., Jacobs, K., 2010a. Option Valuation with Conditional Heteroskedasticity and Nonnormality. The Review of Financial Studies 23, 2139–2183.

Christoffersen, P., Feunou, B., Jacobs, K., Meddahi, N., 2014. The Economic Value of Realized Volatility: Using High-Frequency Returns for Option Valuation. Journal of Financial and Quantitative Analysis 49, 663–697.

Christoffersen, P., Heston, S. L., Jacobs, K., 2006. Option Valuation with Conditional Skewness. Journal of Econometrics 131, 253–284.

Christoffersen, P., Heston, S. L., Jacobs, K., 2009. The Shape and Term Structure of the Index Option Smirk: Why Multifactor Stochastic Volatility Models Work so Well. Management Science 55, 1914–1932.

Christoffersen, P., Heston, S. L., Jacobs, K., 2013. Capturing Option Anomalies with a Variance-Dependent Pricing Kernel. The Review of Financial Studies 26, 1963–2006.

Christoffersen, P., Jacobs, K., 2004a. The Importance of the Loss Function in Option Valuation. Journal of Financial Economics 72, 291–318.

Christoffersen, P., Jacobs, K., 2004b. Which GARCH Model for Option Valuation? Management Science 50, 1204–1221.

Christoffersen, P., Jacobs, K., Mimouni, K., 2010b. Volatility Dynamics for the S&P500: Evidence from Realized Volatility, Daily Returns and Option Prices. The Review of Financial Studies 23, 3141–3189.

Christoffersen, P., Jacobs, K., Ornthanalai, C., 2012. Dynamic Jump Intensities and Risk Premiums: Evidence from S&P500 Returns and Options. Journal of Financial Economics 106, 447–472.

Clark, I. J., 2011. Foreign Exchange Option Pricing: A Practitioner's Guide. John Wiley & Sons.

Cont, R., 2001. Empirical Properties of Asset Returns: Stylized Facts and Statistical Issues. Quantitative Finance 1, 223–236.

Corradi, V., 2000. Reconsidering the Continuous Time Limit of the GARCH (1, 1) Process. Journal of Econometrics 96, 145–153.

Cox, J. C., Ingersoll, J. E., Ross, S. A., 1985. A Theory of the Term Structure of Interest Rates. Econometrica 53, 385.

Cox, J. C., Ross, S. A., 1976. The Valuation of Options for Alternative Stochastic Processes. Journal of Financial Economics 3, 145–166.

Cui, Y., Rollin, S. D. B., Germano, G., 2015. Full and Fast Calibration of the Heston Stochastic Volatility Model. arXiv preprint 1511.08718.

Delbaen, F., Schachermayer, W., 2006. The Mathematics of Arbitrage. Springer Science & Business Media.

Duan, J.-C., 1995. The GARCH Option Pricing Model. Mathematical Finance 5, 13–32.

Duan, J.-C., 1997. Augmented GARCH (p,q) Process and its Diffusion Limit. Journal of Econometrics 79, 97–127.

Duan, J.-C., Ritchken, P., Sun, Z., 2006. Approximating GARCH–Jump Models, Jump–Diffusion Processes, And Option Pricing. Mathematical Finance 16, 21–52.

Duffie, D., Pan, J., Singleton, K., 2000. Transform Analysis and Asset Pricing for Affine Jump–Diffusions. Econometrica 68, 1343–1376.

Dumas, B., Fleming, J., Whaley, R. E., 1998. Implied Volatility Functions: Empirical Tests. The Journal of Finance 53, 2059–2106.

Engle, R. F., 1982. Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. Econometrica: Journal of the Econometric Society pp. 987–1007.

Engle, R. F., Mustafa, C., 1992. Implied ARCH Models from Options Prices. Journal of Econometrics 52, 289–311.

Engle, R. F., Ng, V. K., 1993. Measuring and Testing the Impact of News on Volatility. The Journal of Finance 48, 1749–1778.

Eraker, B., 2004. Do Stock Prices and Volatility Jump? Reconciling Evidence from Spot and Option Prices. The Journal of Finance 59, 1367–1403.

French, K. R., Schwert, G. W., Stambaugh, R. F., 1987. Expected Stock Returns and Volatility. Journal of Financial Economics 19, 3–29.

Gârleanu, N., Pedersen, L. H., 2013. Dynamic Trading with Predictable Returns and Transaction Costs. The Journal of Finance 68, 2309–2340.

Gatheral, J., 2006. The Volatility Surface: A Practitioner's Guide. John Wiley & Sons.

Gilli, M., Schumann, E., 2012. Calibrating Option Pricing Models with Heuristics. In: Brabazon, A., O'Neill, M., Maringer, D. (eds.), *Natural Computing in Computational Finance: Volume 4*, Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 9–37.

Giovanis, E., 2009. Calendar Effects and Seasonality on Returns and Volatility. MPRA Paper No. 64404.

Glosten, L. R., Jagannathan, R., 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, 1779–1801.

Harrison, J. M., Pliska, S. R., 1981. Martingales and Stochastic Integrals in the Theory of Continuous Trading. Stochastic Processes and Their Applications 11, 215–260.

Hentschel, L., 1995. All in the Family Nesting Symmetric and Asymmetric GARCH Models. Journal of Financial Economics 39, 71–104.

Heston, S. L., 1993. A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options. The Review of Financial Studies 6, 327–343.

Heston, S. L., Nandi, S., 2000. A Closed-Form GARCH Option Valuation Model. The Review of Financial Studies 13, 585–625.

Hsieh, K. C., Ritchken, P., 2005. An Empirical Comparison of GARCH Option Pricing Models. Review of Derivatives Research 8, 129–150.

Hull, J., White, A., 1987. The Pricing of Options on Assets with Stochastic Volatilities. The Journal of Finance 42, 281–300.

Itô, K., 1951. On Stochastic Differential Equations. Memoirs of the American Mathematical Society 4, 1–51.

Jackwerth, J. C., 2000. Recovering Risk Aversion from Option Prices and Realized Returns. Review of Financial Studies 13, 433–451.

Kahl, C., Jäckel, P., 2005. Not-So-Complex Logarithms in the Heston Model. Wilmott Magazine 19, 94–103.

Kallsen, J., Taqqu, M. S., 1998. Option Pricing in ARCH–Type Models. Mathematical Finance 8, 13–26.

Kilin, F., 2011. Accelerating the Calibration of Stochastic Volatility Models. Journal of Derivatives 18, 7–16.

Lesmond, D. A., Ogden, J. P., Trzcinka, C. A., 1999. A New Estimate of Transaction Costs. The Review of Financial Studies 12, 1113–1141.

Liu, H., Loewenstein, M., 2002. Optimal Portfolio Selection with Transaction Costs and Finite Horizons. The Review of Financial Studies 15, 805–835.

Lo, A. W., MacKinlay, A. C., 2008. A Non-Random Walk Down Wall Street. Princeton University Press, Princeton.

Lord, R., Kahl, C., 2006. Why the Rotation Count Algorithm Works. Tinbergen Institute Discussion Paper 06-065/2.

Lord, R., Kahl, C., 2010. Complex Logarithms in Heston-Like Models. Mathematical Finance 20, 671–694.

Merton, R. C., 1973. Theory of Rational Option Pricing. The Bell Journal of Economics and Management Science 4, 141.

Moyaert, T., Petitjean, M., 2011. The Performance of Popular Stochastic Volatility Option Pricing Models During the Subprime Crisis. Applied Financial Economics 21, 1059–1068.

Müller, J., Kanniainen, J., Piché, R., 2013. Calibration of GARCH Models Using Concurrent Accelerated Random Search. Applied Mathematics and Computation 221, 522–534.

Munk, C., 2013. Financial Asset Pricing Theory. Oxford University Press, Oxford, New York.

Musiela, M., Rutkowski, M., 2009. Martingale Methods in Financial Modelling, vol. 36 of *Stochastic Modelling and Applied Probability*. Springer, Berlin, second ed.

Nandi, S., 1996. Pricing and Hedging Index Options under Stochastic Volatility: An Empirical Examination. Federal Reserve Bank of Atlanta.

Nelder, J. A., Mead, R., 1965. A Simplex Method for Function Minimization. The Computer Journal 7, 308–313.

Nelson, D. B., 1990. Stationarity and Persistence in the GARCH (1, 1) Model. Econometric Theory 6, 318–334.

Nelson, D. B., 1991. Conditional Heteroskedasticity in Asset Returns: A New Approach. Econometrica 59, 347.

Nielsen, L. T., 1993. Understanding N (d1) and N (d2): Risk Adjusted Probabilities in the Black-Scholes Model. Journal of the French Finance Association 14, 95–106.

Pagan, A. R., Schwert, G. W., 1990. Alternative Models for Conditional Stock Volatility. Journal of Econometrics 45, 267–290.

Pastorello, S., Renault, E., Touzi, N., 2000. Statistical Inference for Random-Variance Option Pricing. Journal of Business & Economic Statistics 18, 358.

Ross, S. A., 1976. Options and Efficiency. The Quarterly Journal of Economics pp. 75–89.

Rouah, F., Vainberg, G., 2007. Option Pricing Models and Volatility Using Excel-VBA (Wiley finance). John Wiley & Sons, Hoboken, N.J.

Rubinstein, M., 1976. The Valuation of Uncertain Income Streams and the Pricing of Options. The Bell Journal of Economics 7, 407.

Schmelzle, M., 2010. Option Pricing Formulae Using Fourier Transform: Theory and Application.

Schöbel, R., Zhu, J., 1999. Stochastic Volatility with an Ornstein–Uhlenbeck Process: An Extension. European Finance Review 3, 23–46.

Schoutens, W., Simons, E., Tistaert, J., 2003. A Perfect Calibration! Now What? The Best of Wilmott p. 281.

Schwert, G. W., 1989. Why Does Stock Market Volatility Change over Time? The Journal of Finance 44, 1115–1153.

Stein, E. M., Stein, J. C., 1991. Stock Price Distributions with Stochastic Volatility: An Analytic Approach. The Review of Financial Studies 4, 727–752.

Storn, R. M., Price, K. V., 1997. Differential Evolution–A Simple and Efficient Heuristic for Global Optimization over Continuous Spaces. Journal of Global Optimization 11, 341–359.

Taylor, S. J., 1986. Modelling Financial Time Series. John Wiley & Sons, Chichester.

Ting, C., 2006. Which Daily Price is Less Noisy? Financial Management 35, 81–95.

Tsiakas, I., 2005. Is Seasonal Heteroscedasticity Real? An International Perspective. Finance Letters 3, 124–132.

Eidesstattliche Erklärung

Ich erkläre hiermit gem. § 5 Abs. 3 PuStO, dass ich die vorstehende Masterarbeit selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.

(Datum) (Unterschrift)