Assessing the Economic Value of Implied Volatility Estimates

Johannes C. Ackermann[*]

February 1, 2018

Part I: The Economic Value of NAGARCH and Markov-Switching Implied Volatility Estimates
Advisor: Massimo Guidolin[†]

Part II: Practical Alternatives: (Ad-hoc) Black-Scholes and the VIX
Advisor: Michael Halling[‡]

M.Sc. Thesis
Department of Finance
Stockholm School of Economics

Abstract

This thesis studies the value of implied volatility estimates for portfolio allocation under the modern portfolio theory (MPT) framework introduced by Markowitz and compares the pricing performances of several common option pricing models. The thesis consists of two parts. The first part compares the NAGARCH framework with the bidirectional Markov switching models (B-MSM) of (Duan et al., 2002) and develops an efficient pricing algorithm for European options under a finite-state model. The variance-implied portfolios are then evaluated on the S&P 500, giving further evidence for using implied volatility estimates for asset allocation. The second part compliments the first by analyzing the value of implied volatility estimates using practical alternatives, namely the Black-Scholes (BS) model, the ad-hoc Black-Scholes (ABS) model and the VIX. The ABS model shows very promising results, outperforming standard BS in form of option pricing and under MPT can create value via volatility timing, beating a buy-and-hold strategy on the S&P 500 in some cases and showing improvements over a VIX induced strategy. However, the simple, practical approaches fall short when compared to the models in part 1, both in form of option pricing as well as MPT asset allocation, suggesting that there is a considerable payoff in estimating implied volatility under these more complex frameworks.

JEL classifications: G11, G12
Keywords: implied volatility, option pricing, modern portfolio theory, GARCH, Markov switching models

[*]40960@student.hhs.se
[†]Bocconi University, massimo.guidolin@unibocconi.it
[‡]Stockholm School of Economics, michael.halling@hhs.se

Disclaimer

This thesis was written in accordance with the regulations of the double degree program in Finance between Bocconi University and Stockholm School of Economics. The first part of this thesis has already been submitted as an independent Master of Science thesis at Bocconi University on November 7, 2017, with the title "Assessing the Economic Value of Modelling Implied Volatility under Regime Switching". The second part was added as an extension to fulfill the requirements of a second Master of Science degree at Stockholm School of Economics. Hence, the first part of this thesis is only to be seen as a revised version of the previously submitted work, edited to remove typos and to improve wording, while the second part contains new research.

Acknowledgements

I thank my friends and family for their moral support during this stressful period, and my advisors, Massimo Guidolin and Michael Halling, for their patience and quick and helpful replies.

Contents

Part I

The Economic Value of NAGARCH and Markov-Switching Implied Volatility Estimates

I.1 Introduction

Since the work of (Markowitz, 1952) and the introduction of the two-factor model for optimal asset allocation, researchers have been driven to find the best stock return and volatility forecasts. Starting with the seminal papers of (Engle, 1982) and (Bollerslev, 1986), generalized autoregressive conditional heteroskedasticity (GARCH) models have long been in the focus of an increasing number of researchers for the analysis and forecasting of volatility. The models generally benefit from the existence of efficient statistical estimation routines and allow for major stylized facts such as volatility clustering. Many extensions of the basic GARCH model have been put forward, with the nonlinear asymmetric GARCH model (NAGARCH) recently gaining increasing popularity by offering to also capture the leverage effect. Studies have shown, however, that GARCH models sometimes appear too smooth and underperform during major unexpected structural changes such as financial crises. Therefore, building on the work of (Hamilton, 1989), research into Markov switching models (MSM) has grown, where the underlying price dynamic is modelled to switch between different volatility regimes at random times, according to a Markov process. Whilst these models thus also allow the variance to influence returns, they originally did not allow returns to affect the variance process. (Duan et al., 2002) has closed this gap by introducing a new class of MSM's that contain a bidirectional feedback mechanism between both processes, therefore matching this property of the GARCH models. This new class of bidirectional Markov switching models (henceforth, B-MSM) is more general than the GARCH framework as it allows for an additional type of shocks to affect the volatility process besides the return innovations, hence allowing the modelling of the aforementioned structural breaks.

Having chosen one of these models, historical return data can be used to estimate the variance over this past period and invest into assets according to the two-factor model. However, there is only so much that can be learned from the past and there is no guarantee that future returns and volatilities will resemble their past. Hence, asset allocation research has moved to forward-looking measures, most notably, the estimation of the volatility implied in observed option prices.

The thesis contributes in several ways. It proposes an efficient estimation procedure for valuing European options in the (Duan et al., 2002) framework and derives option-implied parameter estimates for the bidirectional Markov switching model with 11 states. Then, the thesis shows the value of implied volatility for asset allocation purposes and compares the portfolio performances based on B-MSM with the ones achieved via the NAGARCH model.

The structure of the thesis is as follows. In chapter I.2, the thesis is further placed into the existing literature, providing an overview over many important contributions related to the topics being discussed. Afterwards, in chapter I.3, the frameworks of the NAGARCH and B-MSM models are introduced under the physical and the risk-neutral measure. Here, two special cases of B-MSM are considered, namely the two-state case without a return feedback mechanism in the spirit of the original MSM's, and the limit framework as the number of states approaches infinity. The data set used for the empirical analysis is then described in chapter I.4, including the applied exclusion filters. Then, in chapter I.5 and chapter I.6, the estimation procedures for each model are presented and the results are discussed. The thesis ends with chapter I.7, where the conclusions drawn from the thesis' results are presented.

I.2 Literature Review

Given the nature of this thesis, there are three major research fields the thesis must be placed into: asset allocation, option valuation models and Markov switching models.

I.2.1 Asset Allocation

Modern portfolio theory is based on the two-factor model developed by Markowitz (Markowitz, 1952; Markowitz, 1959). In this model, asset returns are assumed to be random and a portfolio's return is the weighted sum of its individual asset returns, which implies that the expected portfolio return equals a linear combination of the expected returns of all single investments. Markowitz further proposed that the risk of the portfolio should then analogously be expressed by the variance of the portfolio returns. He introduced a notion of efficient portfolios, which are not strictly dominated by any other portfolio in a mean-variance sense. (Markowitz, 1959) and (Tobin, 1958) make the connection between the two-factor model and the expected utility theory of (Von Neumann and Morgenstern, 1945). Tobin also proved an important result, now known as Tobin's separation theorem, which separates the investment decision from the risk aversion of the individual investor. In this sense, everyone invests into the same market portfolio and a risk-free asset, and only the two proportions are determined by the individual risk aversion parameter. Hence, the only inputs the model needs to determine the optimal allocation of capital are the expected value and volatility of the returns of the market portfolio. Given these parameters, the model pins down the optimal portfolio weights depending on the investor's risk aversion.

Since the volatility parameter of the underlying assets is generally not known, one could use the historical variation of asset returns. However, the portfolio allocation problem is concerned with future returns and there is no reason to assume that the past will exactly repeat itself. Hence, an arguably more fitting approach is to use the forward-looking nature of implied volatilities contained in option prices. Since options are also concerned with future returns of the underlying and their market prices already imply a certain volatility of the underlying asset, it seems natural to use this volatility estimate also for these asset allocation purposes. The economic value of option-implied estimates for optimal asset allocations has been shown in similar settings by (Kostakis et al., 2011), who showed that option-implied return distributions can improve portfolio returns compared with historical-based measures, and (Busch et al., 2011), who successfully tested the predictive power of option-implied forecasts versus realized volatility forecasts.

I.2.2 Option Valuation Models

Option valuation models became extensively popular in recent years after the research field first blossomed with the invention of the Black-Scholes model (BS) (Black and Scholes, 1973; Merton, 1973). BS was so ground breaking because it offered a simple, closed-form solution to price European options with the only unobservable parameter to be estimated being the volatility of the underlying asset. Hence, the price of an European option under BS boils down to estimating volatility, with the option price then being given by a simple monotone increasing function. Besides the extreme simplicity, this also shows the major flaw of the model as it assumes that the volatility of the underlying asset does not depend on market movements, an assumption that has been vastly rejected by empirical studies. In fact, by inverting said mapping and retrieving the implied volatilities from observed market prices of European option contracts, researchers have found what is known as volatility smiles and smirks, i.e. non-flat implied volatility surfaces across different moneyness values of option contracts, which is in direct contrast to what is assumed in BS. Given this insufficiency of BS, substantial empirical research has been conducted into alternative option pricing models. These models are usually trying to develop plausible specifications of the processes of the underlying financial variables. By doing so, these models usually imply an associated risk-neutral probability measure that is then used to find the price of a derivative as the discounted value of its expected future payoffs. Besides many other approaches[1], the bulk of modern asset pricing models can reasonably be classified into *stochastic volatility models* (SVMs) and *discrete time volatility models* (DTVMs).

One of the main ideas of SVMs compared to time variant or local volatility models was to describe volatility with a second, stochastic process (next to the already being modelled process for the underlying asset). This finally allowed to incorporate empirical observations such as *volatility clustering*[2] and the *leverage effect*[3]. Another advantage of SVMs is that they typically generate return distributions with fatter left tails and higher peaks than the otherwise commonly assumed normal distribution. Furthermore, (Renault and Touzi, 1996) have shown that a stochastic volatility

[1] Noteworthy mentions include the implied local volatility models advocated by (Dupire, 1994), jump diffusion models after (Merton, 1976) and lattice approach models, which build on the (Cox et al., 1979) binomial tree model.

[2] Volatility clustering stands for the observation that large absolute changes are more likely to be followed by large ones and small ones more likely to be followed by small absolute changes (Mandelbrot, 1963).

[3] The leverage effect describes the phenomenon when negative returns have a stronger impact on volatility than positive returns of the same magnitude (Black, 1976).

model like the Hull-White model (Hull and White, 1987) with independent Brownian motions always results in an implied volatility smile for (vanilla) equity derivatives. The main disadvantage of SVMs is that they introduce a non-tradable source of randomness, making the market incomplete, and pricing and hedging measures no longer unique, which strongly limits their practical applications (Mitra, 2011). Meanwhile, DTVMs mostly build on the autoregressive conditional heteroscedasticity model by (Engle, 1982) and its generalization to (GARCH) models in (Bollerslev, 1986). These frameworks assume that variance is driven by a (discrete) innovation process and past realizations. Empirically, this class of models has been very successful for modelling asset prices, which lead to (Duan, 1995) developing a locally risk neutral measure in the GARCH framework. Economists have extended the basic GARCH model to allow for some of the stylized facts that are captured by the SVM's. Prominent examples are the GJR-GARCH (Glosten et al., 1993) and TGARCH (Zakoian, 1994) models, which incorporate asymmetry in the ARCH process (i.e. volatility clustering), whilst the EGARCH model (Nelson, 1991) incorporates the previously described leverage effect. (Engle and Ng, 1993) proposed the nonlinear asymmetric GARCH model (NAGARCH), which allows for both, volatility clustering and the leverage effect. Many studies have compared different iterations of GARCH settings, proving the empirical strength of GARCH models compared with SVMs (Lehar et al., 2002) and of the NAGARCH setting in particular compared with other GARCH settings (Christoffersen and Jacobs, 2002; Christoffersen and Jacobs, 2004). A major advantage of most DTVM's is that volatility at certain points in time is fully known given the model parameters, thus allowing for unique pricing and hedging measures. On the other hand, DTVMs generally do not allow for closed-form pricing formulas as is typical for SVMs[4]. However, it should be mentioned that (Heston and Nandi, 2000) recently found an analytic expression

[4]For example, see the Hull-White model (Hull and White, 1987), which allows for a closed-form solution for European options even in the case of correlated return and variance processes.

for pricing European options under certain GARCH specifications.

I.2.3 Markov Switching Models

According to (Hamilton and Susmel, 1994), a major drawback of GARCH models is their high persistence of shocks on volatility as empirical research points towards recurrent significant changes in the structure of the markets, resulting in parameter instability. Examples of those studied effects that disturb markets and cause major issues for these models are financial crises (Jeanne and Masson, 2000; Hamilton, 2005) and changes in government policy (Hamilton, 1988; Sims and Zha, 2004). To deal with these changing conditions, economists became increasingly interested in nonlinear, time-varying parameter models that offer the inclusion of such structural breaks.

The Markov switching model (henceforth, MSM), introduced in the seminal work of (Hamilton, 1989), uses several regimes and allows some parameters to switch between them over time, enabling the model to capture more complex data patterns. This switching process is driven by an unobservable latent state variable that follows a Markov process, i.e. each realization of the variable is assumed to only depend on its most immediate past value. It exists a vast amount of literature on using MSMs for financial and economic research, with a comprehensive recent survey provided in (Guidolin, 2011; Guidolin, 2013). Given the empirical success of both MSM and GARCH, it was natural to try and combine both frameworks (Cai, 1994; Gray, 1996; Dueker, 1997; Bollen, 1998; Klaassen, 2002). (Duan et al., 2002) then proposed a new class of MSM that captures the impact of return innovations on the volatility process and thus created the bridge to the GARCH framework. Their models also include the MSM proposed in the seminal work of (Hamilton, 1989) as a limiting case, but in the more general setting allow for a bilateral feedback mechanism between the return and volatility processes to capture more complex empirically observed phenomena. Empirically, MSM have shown very promising results for option pricing. (Bollen et al., 2000) proved that even a basic regime switching model with independent shifts in the mean and variance can dominate several common GARCH specifications in the foreign exchange market. Hence, given the empirical success of MSMs, this thesis attempts to extend the work of (Giudice, 2017), who studied the economic value of NAGARCH-implied volatility estimates, by obtaining the option implied volatility estimates from a B-MSM according to (Duan et al., 2002) instead and by measuring the models' option pricing and portfolio allocation performances[5].

[5]For alternative research on using MSMs for asset allocation purposes see for example (Guidolin and Timmermann, 2007).

I.3 Methodology

This paper considers a stock market index and its corresponding rate of return R_{t+1}, defined by

$$R_{t+1} = \log\left(\frac{S_{t+1}}{S_t}\right), \tag{I.3.1}$$

where S_t denotes the index closing level on day t. Furthermore, a utility maximizing investor is assumed, who is solely concerned with the expected returns and variance of the portfolio, and who is equipped with an information set Ψ_t that contains realizations of all relevant variables up to time t.

I.3.1 Portfolio Optimization Framework

It is assumed that the investor can invest into two assets, a risk-free one such as a 10-year U.S. government bond, and a risky one, say the S&P 500 index (e.g. through buying an ETF such as State Street's *SPY*). Denoting the proportion invested in the S&P 500 and the risk-free asset by x_t^s and x_t^b, and the daily yield of the risk-free asset during the time interval $[t, t+1]$ by r_f[1], the conditional expected value and variance[2] of the portfolio returns are given by

$$\mathrm{E}_t[R_{t+1}^P] = x_t^s \mathrm{E}_t[R_{t+1}^s] + x_t^b r_f \tag{I.3.2}$$

[1]For better readability and conformity with the literature, notation is being simplified to r_f. However, this shall not imply that the rate is taken to be constant over all maturities in the empirical part of this study, but rather that the appropriate rate is chosen at any time and for each maturity.

[2]Henceforth, notation is slightly being abused for better readability, the reader should be aware that $\mathrm{E}_t[R_{t+1}]$ and $\mathrm{V}_t[R_{t+1}]$ always stand for $\mathrm{E}[R_{t+1}|\Psi_t]$ and $\mathrm{V}[R_{t+1}|\Psi_t]$, respectively.

and[3]

$$\mathrm{V}_t[R^P_{t+1}] = (x^s_t)^2 \mathrm{V}_t[R^s_{t+1}]. \tag{I.3.3}$$

The portfolio optimization problem then becomes finding a series of portfolios, which have an efficient trade-off between maximizing (I.3.2) and minimizing (I.3.3). This is formalized as

$$x^*_t = arg \min_{x^s_t, x^b_t} -\kappa \left(x^s_t \mathrm{E}_t[R^s_{t+1}] + x^b_t r_f \right) + \frac{1}{2}(x^s_t)^2 \mathrm{Var}_t[R^s_{t+1}] \quad \text{s.t.} \quad x^s_t + x^b_t = 1, \tag{I.3.4}$$

where κ and $x^s_t + x^b_t = 1$ are the investor's risk aversion and the budget constraint, respectively. A portfolio is called *parametric-efficient* if it is a solution for (I.3.4) for some positive parameter κ. It has been shown (see, for example, (Best, 2010)) that these parametric-optimal portfolios are then given by

$$x^{s*}_t = \frac{\mathrm{E}_t[R^s_{t+1}] - r_f}{\kappa \mathrm{Var}_t[R^s_{t+1}]}, \quad x^{b*}_t = 1 - x^{s*}_t. \tag{I.3.5}$$

Hence, a framework is needed to determine $\mathrm{E}_t[R^s_{t+1}]$ and $\mathrm{Var}_t[R^s_{t+1}]$.[4]

I.3.2 Nonlinear Asymmetric GARCH(1,1)-M Model

The NAGARCH(1,1)-M model is an extension of the basic GARCH(1,1) model[5]:

$$\begin{aligned} R_{t+1} &= r_f + \sigma_{t+1} z_{t+1}, \quad z_{t+1} \overset{iid}{\sim} \mathbb{N}(0,1) \\ \sigma^2_{t+1} &= \omega + \alpha R^2_t + \beta \sigma^2_t. \end{aligned} \tag{I.3.6}$$

[3]Since the second asset is assumed to be free of risk, $\mathrm{V}_t[r_f] = \mathrm{Cov}_t\left[r_f, R^s_{t+1}\right] = 0$. If either the variance or the covariance with the index were positive, the asset would contain intrinsic/extrinsic risk and thus would not be risk-free.

[4]Henceforth, $\mathrm{Var}_t[R_{t+1}] \equiv \sigma^2_{t+1}$.

[5]Empirical evidence usually supports the choice to include only one lag of the ARCH and GARCH processes, as higher orders often do not provide much additional value but increase parameter uncertainty and overfitting risks.

For this model to remain sensible, further conditions are required. Obviously, the variance process needs to remain non-negative and for most practical purposes (in this case, the asset shall be non-risk-free) it is even required to be positive. This can be achieved by setting $\omega > 0$, $\alpha \geq 0$ and $\beta \geq 0$. Furthermore, for the processes to remain stable, the variance process must not diverge over time but needs to remain stationary. It can be shown that this is ensured by enforcing that $\alpha + \beta < 1$, which is known as the stationary condition of the basic GARCH(1,1) model.

The NAGARCH(1,1)-M model extends this framework in several ways. First, the model allows for positive and negative returns to have an asymmetrical impact on the variance process. Second, it extends the ARCH part of the GARCH(1,1) model (αR_t^2) by adding a non-linear bias-adjustment parameter θ. Lastly, the return process is also modelled in a non-linear, heteroscedastic fashion (hence the '-M', 'in-mean', suffix).

For the modelling purposes of this paper, two measures must be introduced: the physical measure $\mathbb{P}$ and the risk-neutral measure $\mathbb{Q}$. The former can be seen as subjective, as it allows for risk aversion to affect asset valuations. As it is human nature, investors are usually risk averse and want to be compensated for risky positions, and hence demand a (unit) risk premium λ for the risk they are taking. Thus, under this measure the fair price of an asset such as an index option would be difficult to pin down since it would depend on each investor's risk aversion parameter. Therefore, the equivalent risk-neutral measure is used for pricing these assets with the property that in an arbitrage-free and complete market there exists a unique price for all assets[6] such that the expected return on all investments equals the risk-free rate.

Both measures are important for the presented asset allocation framework. As will be shown, given the investor's risk aversion the optimal weights under the considered asset allocation framework are uniquely determined by the (unit) risk premium

[6] Second fundamental theorem of asset pricing.

and the index volatility. Whilst the former is clearly part of the physical measure (there is no risk premium in a risk-neutral world), the latter can be estimated both via the physical and the risk-neutral measure, i.e. by obtaining it from historic data or by inferring it from contemporary option prices. The potential advantage of using the implied volatility contained in option prices is that it is based on expectations of future market movements (and thus the market's volatility) and thus is better aligned with the risks the investor is facing.

I.3.2.1 NAGARCH-M under the Physical Measure

The conditional return process R_{t+1} and the conditional variance process σ^2_{t+1} of the NAGARCH(1,1)-M model under the physical measure $\mathbb{P}$ (henceforth, NAGARCH-P) are given by

$$\begin{aligned} R^{\mathbb{P}}_{t+1} &= r_f + \lambda\sigma_{t+1} - \frac{1}{2}\sigma^2_{t+1} + \sigma_{t+1}z_{t+1}, \\ \sigma^2_{t+1} &= \omega + \alpha\sigma^2_t\left(z_t - \theta\right)^2 + \beta\sigma^2_t, \quad z_t \in \mathbb{N}(0,1). \end{aligned} \tag{I.3.7}$$

For the variance σ^2_{t+1} to remain well defined, it is again necessary that $\omega > 0$, $\alpha \geq 0$ and $\beta \geq 0$. The unconstrained parameter θ is included to allow for the *leverage effect* (Black, 1976), where $\theta > 0$ results in negative returns to have a larger impact on the variance process than positive ones of the same magnitude. The persistence of the variance process can be proved to be $\alpha\left(1+\theta^2\right)+\beta$, so that the unconditional variance becomes $\sigma^2 = \frac{\omega}{1-\alpha(1+\theta^2)-\beta}$. Thus, the stationary condition of NAGARCH-P is $\alpha\left(1+\theta^2\right)+\beta < 1$.

The conditional expected return under $\mathbb{P}$ is then given by

$$\mathrm{E}^{\mathbb{P}}_t[R_{t+1}] = r_f + \lambda\sigma_{t+1} - \frac{1}{2}\sigma^2_{t+1}, \tag{I.3.8}$$

which plugged into (I.3.5) leads to

$$x_t^{s*} = \frac{r_f + \lambda\sigma_{s,t} - \frac{1}{2}\sigma_{s,t}^2 - r_f}{\kappa\sigma_{s,t}^2} = \frac{1}{\kappa}\left(\frac{\lambda}{\sigma_{s,t}} - \frac{1}{2}\right), \quad x_t^{b*} = 1 - x_t^{s*} \tag{I.3.9}$$

where $\sigma_{s,t}^2$ follows (I.3.7).

This is the central equation of this thesis, since all future models will be using the same return process, only differentiating themselves in the way they model the variance process. Hence, for the asset allocation part of this thesis, λ will be uniquely determined under the NAGARCH-P framework, whereas the variance estimation will depend upon each framework.

I.3.2.2 NAGARCH-M under the Risk-Neutral Measure

Following (Duan, 1995), a locally risk-neutral valuation relationship (LRNV) exists, consisting of a measure $\mathbb{Q}$ equivalent to $\mathbb{P}$, if either of the following holds:

1. The investor's utility function has constant relative risk aversion and changes to the logarithmic aggregate consumption have conditional normal distributions with constant mean and variance under $\mathbb{P}$.
2. The investor expresses constant absolute risk aversion and changes to the aggregate consumption have conditional normal distributions with constant mean and variance under $\mathbb{P}$.
3. The utility function is linear[7].

Assuming that the conditions for the LRNV are satisfied, the expected return and variance under $\mathbb{Q}$ become

$$\mathrm{E}_t^{\mathbb{Q}}[R_{t+1}] = r_f \tag{I.3.10}$$

[7] Note that the basic assumption on the investor, to be solely concerned with the return and variance of the portfolio, does not necessarily imply that the utility function is quadratic, as long as the returns are expected to be following a normal distribution.

and

$$\mathrm{Var}_t^{\mathbb{Q}}[R_{t+1}] = \mathrm{Var}_t^{\mathbb{P}}[R_{t+1}] = \sigma_{t+1}^2. \tag{I.3.11}$$

This implies that (I.3.7) under $\mathbb{Q}$ is given by

$$\begin{aligned} R_{t+1}^{\mathbb{Q}} &= r_f - \frac{1}{2}\sigma_{t+1}^2 + \sigma_{t+1} z_{t+1}^*, \\ \sigma_{t+1}^2 &= \omega + \alpha\sigma_t^2 \left(z_t^* - \theta^*\right)^2 + \beta\sigma_t^2, \quad z_t^* \in \mathbb{N}(0,1). \end{aligned} \tag{I.3.12}$$

with $z_t^* = z_t + \lambda$ and $\theta^* = \theta + \lambda$. Again, $\omega > 0$, $\alpha \geq 0$ and $\beta \geq 0$ and the stationary condition is $\alpha\left(1 + \theta^{*2}\right) + \beta < 1$, leading to the stationary variance under the risk-neutral measure, $\sigma^2 = \frac{\omega}{1-\alpha(1+\theta^{*2})-\beta}$.

I.3.3 Markov Regime-Switching Model

In this section, the variance is no longer allowed to assume all potential values, as was the case in the NAGARCH framework before, but instead a set of fixed values is imposed, which the variance process is allowed to take at any point. In a Markov switching model (MSM), the variance is usually assumed to remain in a specific volatility regime for a random amount of time before randomly switching over into another regime. Here, the switches between these states occur according to a Markov process, i.e. the probability distribution of moving from one state to another does not depend on the states of the process past the current one. This thesis uses the formulation of (Duan et al., 2002), which implements the switching mechanic via a threshold model and in the most general framework assumes that the regime switching probabilities of the latent variance process are affected by return innovations. This way, a bidirectional feedback mechanism is established between the variance and the return process, which is in line with many empirical observations.

This thesis first considers the general bidirectional switching model (B-MSM) by Duan et al., under the physical measure and the risk-neutral measure. Afterwards, the unidirectional model with two regimes is presented as a special case that allows for closed-form solutions for European option prices, but does not include a feedback mechanism from returns to volatilities. Then, in subsection I.3.3.3 the limit model (B-MSM-Inf) of the bidirectional framework, when the number of regimes approaches infinity, is shown, which ultimately creates the bridge to the NAGARCH framework presented in the previous section as a special case of B-MSM-Inf.

I.3.3.1 Bidirectional-MSM

The return process in the bidirectional MSM with $K < \infty$ states under the physical measure follows

$$R_{t+1}^{(K)} = \ln\left(\frac{S_{t+1}^{(K)}}{S_t^{(K)}}\right) = r_f + \lambda\sigma_{t+1}^{(K)} - \frac{1}{2}\sigma_{t+1}^{(K)2} + \sigma_{t+1}^{(K)} z_{t+1}, \quad z_{t+1} \in \mathbb{N}(0,1), \tag{I.3.13}$$

where the (K) index indicates that the model of the return process depends through the volatility process on the total number of regimes chosen. Of course, at each point in time there will only be one realization of the return process, but under this framework (other than in the NAGARCH model) knowledge of the return realization is not sufficient to pin-down the current volatility estimate. Thus, the return process moving forward is modelled depending on which regime the process might visit and thus, the model changes with the total number of regimes considered. The volatility process with K regimes is driven by a Markov chain that is fully determined by the $K \times K$ transition matrix between volatility states. While in the NAGARCH(1,1) model the volatility process σ_{t+1} was solely subject to the return innovations z_t and its past value σ_t, in the MSM model a second innovation term, ξ_t, is introduced as an unobservable state variable that is independent of the return innovations. The overall impact of z_t and ξ_t on volatility is then determined via the following updating function[8]:

$$F(z_t, \xi_t) = q_1(z_t - \theta)^+ + q_2(z_t - \theta)^- + (1 - q_1 - q_2)|\xi_t|, \tag{I.3.14}$$

where $q_1 \geq 0, q_2 \geq 0$ and $q_1 + q_2 \leq 1$ are weights assigned to the positive and negative realizations of the return innovations z_t, subject to the bias adjustment θ. Hence, similarly to the NAGARCH model before, this model also allows for the previous form of the leverage effect ($\theta > 0$) to create asymmetries in the volatility responses to return innovations. However, this more general setting further amplifies (or dampens) this asymmetry using the weights q_1 and q_2. For instance, the leverage effect can also be seen if $\theta = 0$ and $q_2 > q_1$. Additionally, if the weights q_1 and q_2 are positive, then return innovations are fed back into the volatility process and the structure of (I.3.14) captures volatility clustering, i.e. the phenomenon that large absolute returns are likely be following large absolute returns.

Assuming K volatility levels $\{\delta_1, \delta_2, \ldots, \delta_K\}$ and corresponding threshold values $\{c_0(\delta_i), c_1(\delta_i), \ldots, c_K(\delta_i)\}$ satisfying $c_0(\delta_i) = 0$ and $c_K(\delta_i) = \infty$, the volatility is determined by

$$\sigma_{t+1}^{(K)} = \delta_i \quad \text{if} \quad c_{i-1}\left(\sigma_t^{(K)}\right) \leq F(z_t, \xi_t) < c_i\left(\sigma_t^{(K)}\right) \quad \text{for } i = 1, 2, \ldots, K, \tag{I.3.15}$$

with $F(z_t, \xi_t)$ given by (I.3.14). Regarding the choice of volatility regimes, the following partition condition must be satisfied:

$$\begin{aligned} \delta_1(K) \to 0 \quad \text{and} \quad \delta_K(K) \to \infty \quad \text{as} \quad K \to \infty, \\ \sup_{i \in \{1,2,\ldots,K-1\}} [\delta_{i+1}^2(K) - \delta_i^2(K)] \to 0 \quad \text{as} \quad K \to \infty. \end{aligned} \tag{I.3.16}$$

[8]With the usual notation $(z_t - \theta)^+ \equiv \max(z_t - \theta, 0)$ and $(z_t - \theta)^- \equiv \max(\theta - z_t, 0)$.

The bi-directional K-state MSM under the physical measure $\mathbb{P}$ is then given by

$$\begin{aligned} R^{\mathbb{P}}_{t+1}(K) &= r_f + \lambda\sigma^{(K)}_{t+1} - \frac{1}{2}\sigma^{(K)2}_{t+1} + \sigma^{(K)}_{t+1} z_{t+1}, \\ \sigma^{(K)}_{t+1} = \delta_i \quad \text{if} \quad & c_{i-1}\left(\sigma^{(K)}_t\right) \leq q_1(z_t - \theta)^+ + q_2(z_t - \theta)^- \\ &+ (1 - q_1 - q_2)|\xi_t| < c_i\left(\sigma^{(K)}_t\right), \\ \text{for } i = 1, 2, \ldots, K, \, & t \in \{0, 1, \ldots, T-1\}, 0 \leq q_1, \; 0 \leq q_2, \; q_1 + q_2 \leq 1, \\ & \begin{bmatrix} z_{t+1} \\ \xi_{t+1} \end{bmatrix} |\Psi_t \overset{\mathbb{P}}{\sim} \mathbb{N}\left(\mathbf{0}_{2\times1}, \boldsymbol{I}_{2\times2}\right), \end{aligned} \tag{I.3.17}$$

with $(\mathbf{0}_{2\times1}, \boldsymbol{I}_{2\times2})$ denoting a zero vector and the identity matrix, respectively. However, assuming that $q1$, $q2$ and θ are known, the state transition probabilities still depend on the choice of threshold values $c_i(\sigma^{(K)}_t)$. Hence, following Duan et al., the following structure is imposed:

$$c_i\left(\sigma^{(K)}_t\right) = \sqrt{\max\left(\frac{\frac{1}{2}\left[\delta^2_i(K) + \delta^2_{i+1}(K)\right] - \omega}{\alpha\sigma^{(K)2}_t} - \frac{\beta}{\alpha}, 0\right)} \quad \text{for } i = 1, \ldots, K-1, \tag{I.3.18}$$

with $\omega > 0$, $\alpha > 0, \beta > 0$, $c_0\left(\sigma^{(K)}_t\right) = 0$ and $c_K\left(\sigma^{(K)}_t\right) = \infty$. The additional parameters ω, α, β are introduced in this way, so that it can be shown that the model, given certain assumptions, converges towards the NAGARCH model as the number of regimes approaches infinity (subsection I.3.3.3).

Assuming that the conditions for the existence of an LRNV stated in subsection I.3.2.2 are satisfied, local risk-neutralization leads to

$$
\begin{aligned}
R^{\mathbb{Q}}_{t+1}(K) &= r_f - \frac{1}{2}\sigma^{(K)2}_{t+1} + \sigma^{(K)}_{t+1} z^*_{t+1}, \\
\sigma^{(K)}_{t+1} = \delta_i \quad \text{if} \quad c_{i-1}\left(\sigma^{(K)}_t\right) &\leq \; q_1(z^*_t - \theta^*)^+ + q_2(z^*_t - \theta^*)^- \\
&+ (1 - q_1 - q_2)|\xi^*_t - \nu| < c_i\left(\sigma^{(K)}_t\right) \\
\text{for } i = 1, 2, \ldots, K, t \in \{0, 1, \ldots, T-1\},\ & 0 \leq q_1,\ 0 \leq q_2,\ q_1 + q_2 \leq 1, \\
\begin{bmatrix} z^*_{t+1} \\ \xi^*_{t+1} \end{bmatrix} &\equiv \begin{bmatrix} z_{t+1} + \lambda \\ \xi_{t+1} + \nu \end{bmatrix} |\Psi_t \overset{\mathbb{Q}}{\sim} \mathbb{N}\left(\mathbf{0}_{2\times 1}, \boldsymbol{I}_{2\times 2}\right),
\end{aligned}
\tag{I.3.19}
$$

where ν is a correction parameter similar to λ, which locally risk-neutralizes ξ_t.

I.3.3.2 Two-State Unidirectional MSM

The unidirectional model (U-MSM) follows from (I.3.17) and (I.3.18) when setting $q_1 = q_2 = 0$, i.e. by removing the (potentially asymmetric) impact of return innovations on the state transition probabilities.

U-MSM with two states under the physical measure is then defined as

$$
\begin{aligned}
R^{\mathbb{P}}_{t+1}(K) = r_f + \lambda\sigma^{(K)}_{t+1} - \frac{1}{2}\sigma^{(K)2}_{t+1} &+ \sigma_{t+1} z_{t+1}, \\
\sigma^{(K)}_{t+1} &= \delta_1 \quad \text{if } 0 \leq |\xi_t| < c_1\left(\sigma^{(K)}_t\right), \\
\sigma^{(K)}_{t+1} &= \delta_2 \quad \text{if } |\xi_t| \geq c_1\left(\sigma^{(K)}_t\right), \\
\text{and } & \begin{bmatrix} z_{t+1} \\ \xi_{t+1} \end{bmatrix} |\Psi_t \overset{\mathbb{P}}{\sim} \mathbb{N}\left(\mathbf{0}_{2\times 1}, \boldsymbol{I}_{2\times 2}\right).
\end{aligned}
\tag{I.3.20}
$$

The transition probabilities for the Markov process are then obtained as follows. Denote the transition probabilities for switching from regime i to regime j by $p_{ij} \equiv P^{\mathbb{P}}\left(\sigma_{t+1} = \delta_j | \sigma_t = \delta_i\right)$ for $i = 1, 2$, $j = 1, 2$. The transition probability for remaining in regime 1 is then given by

$$\begin{aligned} p_{11} &\equiv P^{\mathbb{P}}\left(0 \leq |\xi_t| < c_1(\sigma_t) \right| \sigma_t = \delta_1) \\ &= \Phi\left(c_1(\delta_1)\right) - \Phi\left(-c_1(\delta_1)\right), \end{aligned} \tag{I.3.21}$$

where Φ denotes the cumulative standard normal distribution. Obtaining the other probabilities analogously, yields the state transition matrix

$$\begin{bmatrix} p_{11} & p_{12} \\ p_{22} & p_{21} \end{bmatrix} = \begin{bmatrix} \Phi\left(c_1(\delta_1)\right) - \Phi\left(-c_1(\delta_1)\right) & 1 - \Phi\left(c_1(\delta_1)\right) + \Phi\left(-c_1(\delta_1)\right) \\ \Phi\left(c_1(\delta_2)\right) - \Phi\left(-c_1(\delta_2)\right) & 1 - \Phi\left(c_1(\delta_2)\right) + \Phi\left(-c_1(\delta_2)\right) \end{bmatrix}. \tag{I.3.22}$$

By locally risk-neutralizing (I.3.20) in the same way as the bi-directional MSM, the dynamics under the risk-neutral measure are obtained:

$$\begin{aligned} R_{t+1}^{\mathbb{Q}}(K) = \ln\left(\frac{S_{t+1}^{(K)}}{S_t^{(K)}}\right) &= r_f - \frac{1}{2}\sigma_{t+1}^{(K)2} + \sigma_{t+1} z_{t+1}^*, \\ \sigma_{t+1}^{(K)} &= \delta_1 \quad \text{if } 0 \leq |\xi_t^* - \nu| < c_1\left(\sigma_t^{(K)}\right), \\ \sigma_{t+1}^{(K)} &= \delta_2 \quad \text{if } |\xi_t^* - \nu| \geq c_1\left(\sigma_t^{(K)}\right), \\ \begin{bmatrix} z_{t+1}^* \\ \xi_{t+1}^* \end{bmatrix} &|\Psi_t \overset{\mathbb{Q}}{\sim} \mathbb{N}\left(\mathbf{0}_{2\times 1}, \boldsymbol{I}_{2\times 2}\right). \end{aligned} \tag{I.3.23}$$

The transition probability matrix then becomes

$$\begin{bmatrix} \Phi\left(\nu + c_1(\delta_1)\right) - \Phi\left(\nu - c_1(\delta_1)\right) & 1 - \Phi\left(\nu + c_1(\delta_1)\right) + \Phi\left(\nu - c_1(\delta_1)\right) \\ \Phi\left(\nu + c_1(\delta_2)\right) - \Phi\left(\nu - c_1(\delta_2)\right) & 1 - \Phi\left(\nu + c_1(\delta_2)\right) + \Phi\left(\nu - c_1(\delta_2)\right) \end{bmatrix}. \tag{I.3.24}$$

As (Duan et al., 2002) show, the price of an European call option contract with maturity T and strike price X for given volatility regimes δ_1 and δ_2, current stock price

S_t and risk-free rate r_f, conditional on the current volatility being δ_k, is given by

$$
\begin{aligned}
C_0^k(X,T) &= \sum_{j=0}^{T} \gamma_{Tj}^k C_{j0}^k \\
\text{with } C_{j0}^k &= S_0\Phi(d_{1j}) - Xe^{-r_f T}\Phi(d_{2j}), \\
d_{1j} &= \frac{\ln(S_0/X) + Tr_f + \phi_j^2/2}{\phi_j}, \\
d_{2j} &= d_{1j} - \phi_j, \\
\phi_j^2 &= j\delta_1^2 + (T-j)\delta_2^2.
\end{aligned} \tag{I.3.25}
$$

γ_{Tj}^k represents the risk-neutral probability of the process being in regime 1 j times and $T-j$ times in regime 2, after initially being in state k ($k = 1,2$). The option prices under the 2-state unidirectional MSM can thus be seen as weighted BS prices. It should be noted that (I.3.25) implies that neither variance is allowed to equal zero, but since the thesis is restricting variances to remain positive due the nature of the risky asset, this is of no further concern. To obtain γ_{Tj}^k, let $F^1(i)$ be the probability of the first (re)visit to regime 1 happening after i periods under the condition that the process started in regime 1. γ_{Tj}^1 (and γ_{Tj}^2 analogously) is then determined for $j = 0,\ldots,T$ as

$$
\begin{aligned}
\gamma_{mk}^1 &= \sum_{j=1}^{m-k+1} F^1(j)\gamma_{m-j,k-1} && \text{for } m = 1,\ldots,T;\ k = 2,\ldots,m, \\
\text{with } F^1(m) &= p_{12}p_{22}^{m-2}p_{21} && \text{for } m = 2,\ldots,T, \\
\gamma_{m1}^1 &= p_{11}p_{12}p_{22}^{m-2} + (m-2)p_{12}^2p_{21}p_{22}^{m-3} + p_{12}p_{21}p_{22}^{m-2} && \text{for } m = 2,\ldots,T, \\
\gamma_{m0}^1 &= p_{12}p_{22}^{m-1} && \text{for } m = 1,\ldots,T, \\
\text{and } \gamma_{11}^1 &= F^1(1) = p_{11}.
\end{aligned} \tag{I.3.26}
$$

I.3.3.3 B-MSM-Inf

If the return and variance processes are following (I.3.17), with the threshold values determined by (I.3.18), the partition condition being satisfied and the current stock price and volatility known, then the B-MSM-K model, as the number of states approaches infinity, converges *almost surely* in $\mathbb{P}$ over the time interval $[0, T]$ towards[9]:

$$\begin{aligned} R^{\mathbb{P}}_{t+1} &= r_f + \lambda\sigma_{t+1} - \frac{1}{2}\sigma^2_{t+1} + \sigma_{t+1}z_{t+1}, \quad z_{t+1} \in \mathbb{N}(0,1) \\ \sigma^2_{t+1} &= \omega + \beta\sigma^2_t + \alpha\sigma^2_t \left[q_1(z_t-\theta)^+ + q_2(z_t-\theta)^- + (1-q_1-q_2)|\xi_t|\right]^2. \end{aligned} \tag{I.3.27}$$

It is important to recognize that B-MSM-Inf no longer assumes a discrete state space, but rather a continuum of regimes with infinite possible volatility levels. The model is again subject to the usual feasibility conditions, i.e. non-negativity or positivity of ω, α and β, and the usual restrictions on the weights q_1 and q_2. Furthermore, the stationary variance can be computed[10] and again needs to remain below 1. Now, if $q_1 = q_2 = \frac{1}{2}$, i.e. by removing the state innovation term ξ_t and assigning the same weights to the positive and negative bias-adjusted return innovations, the volatility process in (I.3.27) becomes

$$\sigma^2_{t+1} = \omega + \beta\sigma^2_t + \alpha\sigma^2_t \left[\frac{(z_t-\theta)^+ + (z_t-\theta)^-}{2}\right]^2 \quad z_t \in \mathbb{N}(0,1) \tag{I.3.28}$$

Defining $\tilde{\alpha} := \frac{\alpha}{4}$ then leads to the NAGARCH-P specification in (I.3.7), since $(z_t - \theta)^+ + (z_t-\theta)^- = z_t - \theta$.

Locally risk-neutralizing (I.3.27) analogously leads to the B-MSM-Inf formulation

[9] For a summary of the proof provided in (Duan et al., 2002), see section A.1.

[10] The formula for the stationary variance is being omitted at this point due to its lengthy and convoluted expression. For details on its calculation it is referred to (Duan et al., 2002).

under the risk-neutral measure $\mathbb{Q}$:

$$\begin{aligned}
R_{t+1}^{\mathbb{Q}} &= r_f - \frac{1}{2}\sigma_{t+1}^2 + \sigma_{t+1} z_{t+1}^*, \\
\sigma_{t+1}^2 &= \omega + \beta\sigma_t^2 + \alpha\sigma_t^2 \left[q_1 (z_t^* - \theta^*)^+ + q_2 (z_t^* - \theta^*)^- + (1 - q_1 - q_2)|\xi_t^*| \right]^2, \\
&\begin{bmatrix} z_{t+1}^* \\ \xi_{t+1}^* \end{bmatrix} |\Psi_t \overset{\mathbb{Q}}{\sim} \mathbb{N}\left(\mathbf{0}_{2\times 1}, \boldsymbol{I}_{2\times 2} \right),
\end{aligned} \tag{I.3.29}$$

with $\theta^* = \theta + \lambda$. Again, setting $q_1 = q_2 = \frac{1}{2}$ leads to the risk-neutral NAGARCH(1,1)-M (henceforth, NAGARCH-Q) dynamics in (I.3.12).

I.4 Data

I.4.1 Overview

chapter II.3 The empirical analysis is conducted using data on Standard & Poor's 500 (henceforth, S&P 500) European-style index options. The data is obtained from Optionmetrics, which is part of the Wharton Research Data Services of the University of Pennsylvania. As pointed out in (Heston and Nandi, 2000), the market for S&P 500 index options is the second most active index options market in the United States and the largest in terms of open interest in options[1]. The data set spans from January 01, 2010 to July 31, 2014. It has become very common in recent research papers testing option valuation models to use a time frame of approximately three to five years[2], which seems to be an ideal choice balancing results stability with potential overfitting issues, and, as pointed out in (Christoffersen and Jacobs, 2002), it is generally desired to use a fairly long time series when dealing with models relying on highly persistent volatility processes.

The data set is sampled each Wednesday (or on the next available trading day in case of a holiday), in-line with the previously mentioned research papers, resulting in 239 recordings. These recordings are split into four periods, the first week is used as an in-sample period for parameter estimation, and the following three weeks are

[1](Rubinstein, 1994) even considers it to be one of the best markets for testing valuation models for European-style options.

[2]See, for example, (Christoffersen and Jacobs, 2004; Heston and Nandi, 2000; Bakshi et al., 1997).

used for out-of-sample model evaluation. Regarding the price data, the observations consist of the closing best bid and best ask prices of the option contracts, and the closing index price, thus leveling potential time discrepancy problems.
Since many of the stocks pay dividends, the daily cash dividends for the S&P 500 are collected, which are available in the S&P 500 information bulletin. For the risk-free rate proxy, the LIBOR curve is taken and linear interpolation between the data points is performed, with the maturities left of the short-end of the curve being approximated by the closest available LIBOR rate.

I.4.2 Dividend Adjustment

Since European-style options cannot be exercised prior to maturity, the spot stock price must be adjusted for discrete (expected) dividends. Denoting the dividend payment s days from time t with $D_t(s)$ and the corresponding daily yield with r_f, the present value at time t of the dividend payments for an option expiring after τ days is given by

$$PVD_t = \sum_{s=1}^{\tau} e^{-r_f s} D_t(s). \tag{I.4.1}$$

PVD_t is then subtracted from the time-t index spot level and the resulting, dividend-adjusted, value is used for all further analysis. This procedure is done on each day for all option maturities, which means that each day every option might face a different (adjusted) index price depending on the maturities of the options and their (assumed to be known) future dividend payments.

I.4.3 Exclusion Filters

Several exclusion filters are applied. First, options with maturities shorter than six days or longer than 100 days are excluded as these products are often driven by additional market forces that are not part of this study (e.g. liquidity-related biases). Second, to ease potential issues with price discreteness for option valuation, mid prices (i.e. the averages of the bid and ask quotes) below $\$0.30$ are removed from both data sets. Third, also to avoid liquidity-related biases, thinly traded options are excluded, with an arbitrary threshold set at 50 contracts per day. Fourth, following (Dumas et al., 1998), only options with absolute moneyness less than or equal to 0.1 are included in both samples, since deep in- and out-of-the-money options have very small time premia and hence contain little information about the volatility function. Here, absolute (forward) moneyness is defined as $|M_t| := |K/F_t - 1|$, where K and F_t denote the strike price of the option and the forward price of the index, respectively[3].

Fifth, a series of no-arbitrage conditions are imposed, following the work of (Gonçalves and Guidolin, 2006):

1. Maturity Monotonicity ($\tau_2 > \tau_1$):

$$c_t(\tau_2) - c_t(\tau_1) \geq 0 \tag{I.4.2a}$$

$$p_t(\tau_2) - p_t(\tau_1) - K\left[e^{-r_f\tau_1} - e^{-r_f\tau_2}\right] \tag{I.4.2b}$$

[3]As (Häfner, 2004) and (Natenberg, 1994) have pointed out, although the spot moneyness, which uses the spot price instead of the forward price, is often preferred by traders, the forward based measure has a more favourable theoretic behaviour. Both authors point towards $\log(K/F_t)$ as the preferred definition for moneyness and $K/F_t - 1$ is then a simple approximation from a first order Taylor series expansion, offering the easier interpretations of moneyness as a proportion of the forward price of the underlying index. The definition is therefore also being used in (Heston and Nandi, 2000) and (Dumas et al., 1998).

2. Reverse Strike Monotonicity ($K_1 > K_2$):

$$[K_1 - K_2]\, e^{-r_f \tau} - c_t(K_2) + c_t(K_1) \geq 0 \tag{I.4.3a}$$

$$p_t(K_1) - p_t(K_2) - [K_1 - K_2]\, e^{-r_f \tau} \geq 0 \tag{I.4.3b}$$

3. Box spreads:

$$[p_t(K_1) - c_t(K_1)] - [p_t(K_2) - c_t(K_2)] - [K_1 - K_2]\, e^{-r_f \tau} \geq 0 \tag{I.4.4a}$$

$$[K_1 - K_2]\, e^{-r_f \tau} - [p_t(K_1) - c_t(K_1)] + [p_t(K_2) - c_t(K_2)] \geq 0 \tag{I.4.4b}$$

4. Maturity spreads:

$$[p_t(\tau_1) - c_t(\tau_1)] - [p_t(\tau_2) - c_t(\tau_2)] - K\, [e^{-r_f \tau_1} - e^{-r_f \tau_2}] \geq 0 \tag{I.4.5a}$$

$$[p_t(\tau_2) - c_t(\tau_2)] - [p_t(\tau_1) - c_t(\tau_1)] + K\, [e^{-r_f \tau_1} - e^{-r_f \tau_2}] \geq 0 \tag{I.4.5b}$$

$p_t(\tau)$ and $c_t(\tau)$ denote the prices of put and call options at day t with maturity in τ days and strike price K. Of course, in practical applications, exploiting these potential arbitrage violations is subject to various market frictions, with the most notably one being transaction costs. To account for the latter, each no-arbitrage condition above is being evaluated using the available best bid and ask prices corresponding to the short and long positions taken (hence the bid-ask spread is used as a proxy for the total transaction costs, see subsection I.5.3.2). This, however, also means that the inverse of each condition must be checked as well, to ensure that there is no arbitrage opportunity in neither the bid nor the ask prices of the given option. To illustrate this, (I.4.2a) is being tested as:

$$c_t^{ask}(\tau_2) - c_t^{bid}(\tau_1) \geq 0 \tag{I.4.6a}$$

$$c_t^{ask}(\tau_1) - c_t^{bid}(\tau_2) \leq 0. \tag{I.4.6b}$$

To further ease the immense computational burden of the (weekly) non-linear optimization routines later on, and to remain consistent with the aforementioned papers, the thesis reduces its scope to call option prices, although the analysis can also be performed equivalently for put options. This results in a total of 18,280 data points.

I.4.4 Summary

Here, a short overview of some key characteristics of the data set is provided. The call contracts are classified by moneyness and days to maturity (DTM). In particular, the option contracts are split into *short* ($6 \leq DTM \leq 50$) and *long* ($51 \leq DTM \leq 100$) term contracts. Since the parameter estimation will be performed every four weeks, Table B.1 shows the number of call contracts by maturity and moneyness for the in-sample week (when estimation is performed) and the following three weeks (where out-of-sample performance is measured). Note that due to the exclusion filters applied, deep-in-the-money[4] and deep-out-of-the-money contracts are excluded. The data set is thus bulked around at-the-money contracts, with a slight concentration of out-of-the-money call contracts. Regarding maturity, the data set is concentrated around short term contracts, with approximately twice as many short term contracts as contracts with long term maturities. Meanwhile, Table B.2 shows the corresponding call prices, which range between approximately $\$1$ and $\$160$, with the long term

[4] Note that by the definition of maturity used, $M_t = K/F_t - 1$, a call option is in-the-money if $M_t < 0$.

contracts being higher priced due to higher potential gains by limited downside. Also, in-the-money options are higher priced than out-of-the-money options due to a higher payoff for the same terminal index level.

The S&P 500 index during the time period is presented in Figure B.1, showing that the S&P 500 steadily grows after a quite tumultuous start as the U.S. recovered from the aftermaths of the financial crisis in 2008-2009. Figure B.2 and Figure B.3 show two important characteristics that are fundamental for the modelling approach taken in this thesis: First, the returns series appears random with little predictability to be detected, whereas the absolute return series (as a proxy for the unobservable daily volatility) seems less like a white noise process. This observation is supported by a formal Box-Pierce test, rejecting the independence hypothesis of the latter series even at the lowest common significance levels, hence the decision for modelling the second moment instead of the first. Additionally, the figures show clear volatility clustering, where large absolute returns are followed by large absolute returns, which motivates the use of non-linear asymmetric models.

I.5 Parameter Estimation and Performance Measurement

I.5.1 NAGARCH-M

I.5.1.1 Maximum-Likelihood Estimation

Given the historical nature of the physical measure, the parameters of (I.3.7) can conveniently be estimated via maximum likelihood estimation (MLE). Under MLE, focusing on the past of the time series, the return innovation z_t in (I.3.7) can be backed out via $z_t = \frac{R_{t+1} - r_f + \frac{1}{2}\sigma_{t+1}^2}{\sigma_{t+1}} - \lambda$. Thus, given the assumed Gaussian distribution of z_t, the parameters are obtained by maximizing the following log-likelihood function for the observed return time series:

$$\log L = -\frac{(T-1)\log 2\pi}{2} - \frac{\sum_{t=2}^{T} \log \sigma_t^2}{2} - \frac{\sum_{t=2}^{T} \left(R_{t+1} - r_f - \lambda\sigma_{t+1} + \frac{1}{2}\sigma_t^2\right)^2}{2\sigma_{t+1}^2}, \quad \text{(I.5.1)}$$

subject to the feasibility and stationarity conditions presented in subsection I.3.2.1. Particularly, this leads to an estimate of the unit risk premium λ, which is essential for the portfolio allocation purposes considered in the second part of this thesis. For a better maximum likelihood estimate, the data set is extended to daily observations (spanning the same period) for this particular estimation, subject to the same data filters described in chapter I.4.

I.5.1.2 Randomized Quasi Monte Carlo

In order to use the forward-looking information hidden in the implied volatility of option prices, NAGARCH-Q is estimated by fitting the model to observed option prices. Since the multi-period distribution of the GARCH process is unknown, the required index level at maturity, S_T, needs to be simulated via Monte Carlo methods. The procedure proposed in this thesis first uses the physical parameters estimates obtained via MLE and transforms them into their risk-neutral counterparts (see subsection I.3.2.2). Then, to obtain a start value for the risk-neutral variance process at day t, an variance updating rule is used. (Christoffersen and Jacobs, 2004) propose to link the volatility between estimation days via

$$\sigma_{t+1}^2 = \omega + \alpha\sigma_t^2 \left(\left[\left(R_t - r_f + \frac{1}{2}\sigma_t^2 \right) / \sigma_t \right] - \theta^* \right)^2 . \tag{I.5.2}$$

Following (Christoffersen and Jacobs, 2004), the thesis uses the 250 daily returns prior to the new estimation date and the obtained parameter starting values to update the initial volatility estimate.

Given this start value, a large number of quasi random variables are drawn from a 20-dimensional Sobol' low discrepancy sequence (Sobol, 1967). The main idea of using Monte Carlo methods for option pricing is the discrete simulation of a given distribution with each path corresponding to a certain realizations of the distribution approximating variable. The problem with using pseudo random variables lies in the infeasibility of drawing infinite many times and thus the issue of always only facing a certain spectrum of the distribution. Instead, quasi random variables do not attempt to simulate a random process, but take deterministic (often equidistant) values that best match the theoretic distribution and therefore often represent a better image of the distributional paths. In fact, many studies have shown that

quasi Monte Carlo methods converge faster than their pseudo counterparts and often show lower pricing errors for option valuation models. In order to still resemble the random characteristic of the return innovations (and to further reduce the variance among simulation estimates) the iterates are additionally scrambled using the procedures proposed by (Owen, 1998) and (Faure and Tezuka, 2002).
The thus defined randomized quasi Monte Carlo procedure draws 600 observations from each dimension, but drops each first 100 iterates as a burn-in sample, making for a total of $MC = 10000$ simulation paths for each trading day until the maturity of the option. Given this shock matrix, the corresponding $\hat{\sigma}^2_{t+1}$ and $\hat{R}_{t+1}$ values are obtained using (I.3.12) and the simulated path price at maturity is calculated as

$$\hat{S}_i(T) = S_0 e^{\left(\sum_{j=1}^{T} \hat{R}_{i,t+1}\right)} \quad i = 1, 2, \ldots, MC. \tag{I.5.3}$$

Unfortunately, as pointed out in (Duan and Simonato, 1998), a known issue of the standard Monte Carlo approach is that this simulated price may violate rational option pricing bounds along its path, making the price estimate non-sensible. In particular, it is required that the discounted asset price at any point in time is a $\mathbb{Q}$-martingale, i.e.

$$\mathrm{E}^{\mathbb{Q}}\left[e^{-r_f t} S(t) | \Psi_\tau\right] = e^{-r_f \tau} S(\tau) \quad \forall t \geq \tau \geq 0. \tag{I.5.4}$$

This cannot be ensured along all simulation paths using standard Monte Carlo, which in turn may lead to a violation of the non-arbitrage condition

$$C_0(t) > \left(S_0 - K e^{-r_f t}\right)^+. \tag{I.5.5}$$

To solve this problem, (Duan and Simonato, 1998) propose a correction to the standard Monte Carlo simulation procedure, called the empirical martingale simulation (EMS). The method is formalized as follows: Given the terminal stock prices in (I.5.3), the EMS prices at maturity, $S^*_{i,T}$, are defined as

$$S^*_i(T) = S_0 \frac{Z_i(T)}{Z_0(T)}, \tag{I.5.6}$$

with

$$\begin{aligned} Z_i(t) &= S^*_i(t-1)\frac{S_i(t)}{S_i(t-1)} \\ Z_0(t) &= \frac{1}{n}e^{-r_f t}\sum_{i=1}^{MC} Z_i(t), \end{aligned} \tag{I.5.7}$$

for $i = 1, \ldots, MC$, $t = 1, \ldots, T$. The price of a call contract at day t under the NAGARCH-Q model is then calculated as

$$C_t^{NG} = e^{-r_f T}\frac{1}{MC}\sum_{i=1}^{MC}\left(S_i(T)^* - K\right)^+, \tag{I.5.8}$$

satisfying $C_t^{NG} > (S^*_0(t) - Ke^{-r_f t})^+ = (S_0 - Ke^{-r_f t})^+$. The authors show that using EMS also reduces the pricing error of the Monte Carlo estimate and can thus also be seen as a variance reduction technique[1]. The parameters are then estimated using a cross-section of available option contracts at day t via non-linear least squares by minimizing[2]

$$MSE_t = \frac{1}{n}\sum_{j=1}^{n}\left(C_{j,t}^{NG} - C_{j,t}^{mkt}\right)^2 \tag{I.5.9}$$

[1]Other variance reduction techniques such as a control variate technique (controlling for the underlying index level, whose expected value is known under the risk-neutral measure) and an antithetic variate technique were implemented as well, but could not improve the RQMC procedure significantly, which is not untypical for quasi Monte Carlo procedures.

[2]Several studies (Christoffersen and Jacobs, 2003; Bams et al., 2004; Renault, 1997) have shown the importance of the loss function and have frequently identified the mean square error function as the best available choice.

where C_t^{mkt} denotes the market (mid) prices of the cross-section of calls at day t, subject to the feasibility and stationarity conditions described in subsection I.3.2.2.

I.5.2 The Markov Switching Model

I.5.2.1 B-MSM-K: Lattice Approach

To estimate the implied parameters of the bidirectional MSM with K states from option prices, a multinomial lattice approach is used, which establishes a discrete Markov chain approximation that converges towards B-MSM as the number of time steps increases. One of the major drawbacks of standard lattice approaches for option valuation under heteroskedasticity is that due to the inherent path dependence in these models the resulting tree is generally not recombining. As in the classic Binomial option pricing modelling approach, jump sizes are usually uniquely determined by the volatility of the underlying asset. Thus, whilst the jump size under the standard log-normal model (e.g. BS) is constant over time, it varies under heteroscedastic models.

This is a major concern for the goals of this thesis, since non-recombining trees imply an exploding number of nodes to be calculated. For example, the total number of nodes in a non-recombining daily binomial lattice for an option that expires in 100 days is $2^{100} - 1$, which is not computational feasible for the purposes of this paper, given that there are 60 estimation days with a total of approximately 4500 in-sample option contracts and each pricing algorithm needed to be repeated several times during optimization routines.

Whilst the following algorithm originally proposed by (Ritchken and Trevor, 1999) to price options under a NAGARCH model, and adapted in (Duan et al., 2002) for

Markov switching models, is highly efficient and allows for a form of recombining lattice, the approach is thought for pricing single European and American options given fixed model parameters. Since the thesis is only considering European options, but wants to infer the model parameters from several cross-sections of option prices, a few simplifying assumptions are taken, which allows to derive the option prices directly from the modelled terminal stock prices instead of using the usual backward recursion. The model to be estimated is given by the return and variance processes under the risk-neutral measure (I.3.19), with the threshold values set according to (I.3.18).

For a given number of regimes, K, the model parameters to be estimated are $\omega, \alpha, \beta, \theta^*$, q_1, q_2 and $\bar{\sigma}_{t,T}$, where $\bar{\sigma}_{t,T}$ defines the 'level' of the volatility regimes over the time interval $[t, T]$ (see Equation I.5.11). First, the K volatility regimes must be set according to the partition condition. Given an estimate for $\bar{\sigma}_{t,T}$, the thesis proposes the following dynamic, in the spirit of the one used in (Duan et al., 2002):

$$\delta_i^2 = L(K) + (i-1)\frac{U(K) - L(K)}{K-1} \quad \text{for } i = 1, \ldots, K \tag{I.5.10}$$

where

$$\begin{aligned} L(K) &= \bar{\sigma}_{t,T}^2 \left(1 - \frac{\sqrt{K-1}}{l}\right)^+ \\ U(K) &= \bar{\sigma}_{t,T}^2 \left(1 + \frac{\sqrt{K-1}}{l}\right) \end{aligned} \tag{I.5.11}$$

for some positive integer l, which determines how far adjacent variance regimes are spread out. In other words, the partition is centered around the estimated variance scale parameter $\bar{\sigma}_{t,T}$, with the allowed variation limited by K and l.

Next, the order of the lattice approach must be set, i.e. how many different paths are considered at each node of the tree. Assuming an equal number m of downside and upside scenarios and a single horizontal movement of the stock price, each node is followed by $2m+1$ branches. In fact, let (y_t^a, σ_{t+1}) be a discrete random process approximating (y_t, σ_{t+1}) with $y_0^a = y_0 = \ln S_0$. Then, given $(y_t^a, \sigma_{t+1} = \delta_k)$, the multinomial process approximating the conditional normal distribution follows:

$$(y_{t+1}^a | \sigma_{t+1} = \delta_k) = y_t^a + j\eta_k\gamma \tag{I.5.12}$$

for $j = 0, \pm 1, \pm 2, \ldots, \pm m$. Here, γ is the jump size (i.e. the gap between adjacent future nodes) and η_k is the smallest positive integer that guarantees the first two moments of the approximating process to match the theoretical conditional moments $\mathrm{E}_t^{\mathbb{Q}}[y_{t+1}|\sigma_{t+1} = \delta_k] = y_t + r_f - \delta_k^2/2$ and $\mathrm{V}_t[y_{t+1}|\sigma_{t+1} = \delta_k] = \delta_k^2$, whilst also ensuring that the path probabilities remain well-defined (i.e. all probabilities fall into $[0,1]$). To obtain the probabilities for a $j = 0, \pm 1, \ldots, \pm m$-move of the log stock price, (Ritchken and Trevor, 1999) divide each day into m homoscedastic intervals, with the variance then being updated at the end of the day. The path probabilities over the full day are then given by the following trinomial dynamic:

$$P\left(y_{t+1}^a = y_t^a + j\eta_k\gamma | \sigma_{t+1} = \delta_k\right) = P(j;k) \quad j = 0, \pm 1, \pm 2, \ldots, \pm m, \tag{I.5.13}$$

where

$$P(j;k) = \sum_{k_u,\, k_m,\, k_d} \binom{m}{k_u\ k_m\ k_d} p_u^{k_u} p_m^{k_m} p_d^{k_d} \tag{I.5.14}$$

with k_u, k_m, $k_d \geq 0$ such that $m = k_u + k_m + k_d$ and $j = k_u - k_d$, and the three probabilities defined as

$$\begin{aligned} p_u &= \frac{\delta_k^2}{2\eta_k^2\gamma^2} + \frac{(r_f - \delta_k^2/2)\sqrt{1/m}}{2\eta_k\gamma} \\ p_m &= 1 - \frac{\delta_k^2}{\eta_k^2\gamma^2} \\ p_d &= \frac{\delta_k^2}{2\eta_k^2\gamma^2} - \frac{(r_f - \delta_k^2/2)\sqrt{1/m}}{2\eta_k\gamma} \end{aligned} \tag{I.5.15}$$

According to (Ritchken and Trevor, 1999), a way of setting η_k to satisfy these two conditions is to let it be determined by

$$(\eta_k - 1) < \frac{\sqrt{\delta_k}}{\gamma} \leq \eta_k \tag{I.5.16}$$

for a given set of variance regimes δ_k and jump size parameter γ.

It is important to note that for the class of models considered in (Ritchken and Trevor, 1999) the path transition probabilities are time variant as they depend on the volatility at time t. However, as already indicated by the substitution of σ_t with δ_k, under B-MSM there are fixed volatility levels that can be attained on each day. Therefore, the path probabilities conditional on being in a certain volatility regime remain the same over time, allowing for the computation of a constant $m \times K$ matrix $\mathrm{P}(j,k)_{j=0,\pm1,\pm2,\ldots\pm m,\ k=1,2,\ldots,K}$, containing the probability for all price innovations j and volatility regimes δ_k.

Using this forward procedure allows to simulate all log stock prices of the next day. However, whilst some volatility regimes induce 'single-jump' prices with $\eta_k = 1$, some volatility regimes might require 'double-jump' prices with $\eta_k = 2$ for the first two return moments to match the assumed ones, and so on. Since it cannot be known which volatility regime the price process is coming from, all potential price paths, i.e. single and multiple price jumps, need to be considered. To circumvent this computational issue, the thesis proposes the following way of determining the jump size parameter γ:

$$\gamma = \sqrt{\bar{\sigma}^2_{t,T}\left(1 + \frac{\sqrt{K}}{\ell}\right)}. \tag{I.5.17}$$

This contributes in two ways. First, it is a desired property that the jump sizes of the lattice depend on the variance of the underlying asset, and secondly, this specific setting ensures that γ is sufficiently large so that for each modelled variance δ_k sensible probabilities can be found for all $2m+1$ 'single-jump' log prices to match the conditional mean and variances of the conditional normal distribution being approximated. In other words, it allows for $\eta_k \equiv 1\ \forall k$. This implies that the approximate log stock value on day t at 'level' j can be easily computed as

$$y^{a(m)}(t, j) = y_0 + j\gamma \quad \text{for } j = 0, \pm 1, \pm 2, \ldots, \pm tm. \tag{I.5.18}$$

for $t = 1, \ldots, T$, where T denotes the maturity of the option contract.

Now, the option price at each node is given by the discounted expected value of all $2m + 1$ subsequent nodes. Unfortunately, while the tree contains a single stock price at any node, the volatility regime at these nodes cannot be known and thus it is necessary to carry K option prices at each node and the expected value also depends on the transition probabilities between the volatility regimes. At maturity of the option contract though, each option price equals the payoff of the option contract and thus only depends on the approximate log stock price and not the volatility regime at that time. Hence, the K-vectors at all terminal points of the tree contain K equal values.

As illustrated in subsection I.3.3.2 for the unidirectional model with two states, the transition probabilities depend on the threshold values[3]. However, in the bidirectional model, the transition probabilities also depend on q_1, q_2, θ^* and the return innovation z_t. Since by construction it is ensured that the first two conditional moments are matched, the normalized innovations in the lattice, conditional on the path innovation $j = 0, \pm 1, \pm 2, \ldots, \pm m$ and regime $k = 1, 2, \ldots, K$, are expressed as

$$z^*(j, k) = \frac{j\gamma - (r_f - \delta_k^2/2)}{\delta_k}. \tag{I.5.19}$$

Clearly, as m approaches infinity, $z^*(j, k)$ converges *in distribution* towards a standard normal random variable. Let $\pi_{k,p}(j)$ denote the state transition probability

[3]For simplification and following (Duan et al., 2002), it is assumed that $\nu = 0$.

for switching from regime k to regime p given the standardized return innovation $z^*(j,k) = \frac{j\gamma-(r_f-\delta_k^2/2)}{\delta_k}$. Since a Gaussian distribution is assumed for the state variable ξ_t, the transition probabilities are calculated analogue to the reduced case in subsection I.3.3.2 as

$$\begin{aligned}\pi_{k,p}(j) \equiv P^{\mathbb{Q}}\big(c_{p-1}(\delta_k) \leq q_1(z_t^*(j,k)-\theta^*)^+ + q_2(z_t^*(j,k)-\theta^*)^- \\ + (1-q_1-q_2)|\xi_t^*| \; < c_p(\delta_k)\big) \\ = \Phi(\Upsilon_p(j,k)) - \Phi(-\Upsilon_p(j,k)) - \Phi(\Upsilon_{p-1}(j,k)) + \Phi(-\Upsilon_{p-1}(j,k)),\end{aligned} \tag{I.5.20}$$

where Φ again represents the cumulative standard normal distribution and

$$\Upsilon_i(j,k) \equiv \frac{c_i(\delta_k) - q_1(z_t^*(j,k)-\theta^*)^+ - q_2(z_t^*(j,k)-\theta^*)^-}{1-q_1-q_2} \tag{I.5.21}$$

for $i = p-1, p$.

Extending this notion to all regime pairs, a $K \times K$ state transition probability matrix $\Pi(j)$ can be derived. Hence, given the modelled call price on day t at price level j conditional on the current volatility regime being δ_k, the probability associated with the call price on day $t+1$ at price level $j+2$ and volatility regime p, is given by $\pi_{k,p}(2)P(2,k)$. Since this probability is time invariant, the matrices can be combined. Let the j resulting $K \times K$ matrices be represented by Θ_j, with $\Theta_j(k,p) = P(j,k)\pi_{k,p}(j)$.

For simplification, and to further ease the still significant computational burden, $m = 1$ is assumed, which reduces the multinomial lattice to a trinomial one and the number of (one-period) transition probability matrices to 3. Note that by using this algorithm with these proposed simplifications, the number of potential nodes has dropped significantly compared to the initially contemplated non-recombining lattice. In fact, the number of nodes on each day no longer grows exponentially with T, but instead in a linear manner. This allows to conveniently calculate the number of paths over several periods to a certain level j in a similar manner as before. Extending the previous notion and assuming that the volatility today is δ_k, the transition probability matrix of arriving at level h in T days is given by

$$\Theta_h^{(T)} = \sum_{k_u,\, k_m,\, k_d} \begin{pmatrix} T \\ k_u\; k_m\; k_d \end{pmatrix} \Theta_1^{(k_u)} \Theta_0^{(k_m)} \Theta_{-1}^{(k_d)}, \tag{I.5.22}$$

for $h = 0, \pm 1, \pm 2, \ldots, \pm T,\ h = k_u - k_d$ and $k_u, k_m, k_d \in [0, T]$. $\Theta_i^{(m)}$ represents the m-th power of the i-th matrix (or in other words, the product of m-times multiplying the i-th matrix with itself). The elements of the $2T+1$ resulting matrices contain the probabilities associated with the terminal state call prices T periods from today conditional on today's volatility regime. Finally, the call price today (at day t), conditional on the volatility being δ_k, is obtained as

$$C_t^a(k) = e^{-r_f T} \sum_{p=1}^{K} \sum_{h=-T}^{T} \left(e^{y_0 + h\gamma} - K\right)^+ \Theta_h^{(T)}(k, p). \tag{I.5.23}$$

The objective is then to minimize the average of the state dependent mean square errors with the observed market prices of the option contracts on day t:

$$\$MSE = \frac{1}{nK} \sum_{k=1}^{K} \sum_{i=1}^{n} (C_{t,i}^a(k) - C_{t,i}^{mkt})^2. \tag{I.5.24}$$

I.5.2.2 U-MSM-2: Analytical Solution

In the case of the unidirectional two-state Markov switching model, closed-form solutions are available in (I.3.25) and (I.3.26). However, the volatility innovation ξ_t is still present and the volatility regimes are still not identifiable. Therefore, following the suggestion of (Duan et al., 2002), the model is reparameterized with the parameters to be estimated being the volatility regimes, δ_1 and δ_2, and the transition probabilities, p_{11} and p_{22}. Second, given numerically optimized start values for these four parameters, state option prices are calculated using the formulas for both potential current volatility regimes. This setting is then numerically optimized with the objective being to minimize the sum of mean square pricing errors with respect to the observed market prices of the call contracts.

I.5.2.3 B-MSM-Inf: Monte Carlo

Since the limit bidirectional model converges to a framework with a continuum of volatility regimes, the estimation can be done via Monte Carlo. Hence, the procedure is the same as the one described in subsection I.5.1.2, with the sole exception that here a second, independently drawn, set of quasi random variables is generated to model the state innovations ξ. There are six parameters to be estimated, namely ω, α, β, θ^*, q_1 and q_2. Hence, other than in (Duan et al., 2002), the weights q_1 and q_2 in the B-MSM-K and B-MSM-Inf frameworks are not restricted to be equal, offering to capture additional asymmetric impacts of different return innovations on volatility. The start values, including the initial volatility σ_t, are obtained via a numerical search by drawing random numbers from the feasible region of each parameter and comparing the resulting mean square errors for the first estimation period. Subsequent estimations use the previous estimates as start values[4]. The optimization objective is again to minimize the mean square errors with respect to the observed market prices of the call contracts.

[4]Of course, since this optimization routine is highly non-linear, a unique optimum may not exist or cannot be found in a sensible amount of time.

I.5.3 Performance Measurement and Trading Costs

Portfolio re-balancing is assumed to happen monthly, in-line with the parameter estimation. In academic literature, this is often rather done on a daily basis, since this better matches some theoretic model characteristic (e.g. constant risk-free rate), and provides better results before costs. However, daily re-balancing is often not very practical due to trading costs and other restrictions, which is the reason a monthly cycle is chosen. Following the works of (Fleming et al., 2001), (Kolusheva, 2008) and (Giudice, 2017), the performances of the portfolio allocation schemes based on the different variance models are subject to trading costs and are measured via two performance criteria.

I.5.3.1 Performance Measures

Let $\hat{\mu}$ be the sample mean of the model returns of a given allocation scheme and $\hat{\sigma}^2$ the corresponding sample variance. The first measure taken is the classical Sharpe Ratio (SR). The SR for a given strategy j is defined as the mean of the generated excess returns divided by their sample standard deviation, i.e.

$$\hat{SR}_j = \frac{\hat{\mu}_j - r_f}{\hat{\sigma}_j}. \tag{I.5.25}$$

Second, for each scheme a Certainty Equivalent Return (CER) is calculated, i.e. the rate of return that makes the investor indifferent between accepting CER and investing following the scheme in question. CER is commonly formalized as

$$CER_j = \hat{\mu}_j - \frac{\kappa}{2}\hat{\sigma}_j{}^2. \tag{I.5.26}$$

I.5.3.2 Trading Costs

Trading costs are usually split into *explicit* and *implicit* costs. Explicit costs are the realized costs of trading, such as broker commissions and taxes. Indirect costs are related to the optimal execution of the trade, including the bid-ask spread and the price impact of the order (e.g. if a large order needs to be split in several parts, additional implicit costs may occur if the price of the asset changes between the execution of each suborder). Since taxes are country dependent and the price impact of execution may depend on the broker, the technology, the time of the day and many other factors, this thesis will approximate the explicit and implicit trading costs by their two main components, the broker commission and the bid-ask spread. As shown in many studies, transaction costs have generally declined over the years, with average broker commissions on big U.S. equity stocks being more than halved between 1982 and 1992 (Stoll, 1995) to approximately 24 basis points of trade value and further decreased to less than 15 basis points in 2000 (Domowitz et al., 2001). (Keim and Madhavan, 1998) attribute this decline to more competitive market environments and the increased use of low-cost electronic crossing networks by institutional traders. (Domowitz et al., 2001) estimate the combined (one-way) explicit and implicit transaction costs for the NYSE in 2000 at slightly less than 30 basis points of trade value. Hence, since technology innovations and algorithmic execution have further spiked in recent years, but still considering the existence of a necessary fundamental lower bound, total transaction costs of ten basis points of trade value for a stock index such as the S&P 500 nowadays seems appropriate.
Furthermore, as in (Giudice, 2017), leveraged investment in the stock index (by borrowing at the risk-free rate, e.g. by short-selling government bonds) is permitted, but capped at 200% of equity, i.e. $x_t^b \in [-1, 1]$ and $x_t^s \in [0, 2]$. However, in this case, additional borrowing costs occur and are estimated at two basis points of borrowed value. As in other studies on transaction costs such as (Marquering and Verbeek,

1999), it is assumed that there are no additional trading costs associated with the risk-free asset.

Re-balancing is assumed to happen at the beginning of a trading day. Hence, given the previously set portfolio allocation to the risky asset, x_t^s, the effective fraction of the portfolio allocated to the risky asset at the start of period $t+1$ (or the end of period t) is

$$\hat{x}_{t+1}^s = \frac{x_t^s e^{R_{t+1}^s}}{x_t^s e^{R_{t+1}^s} + x_t^b e^{r_f}}, \tag{I.5.27}$$

where R_{t+1}^s and r_f are the returns of the risky and risk-free assets over $[t, t+1]$, respectively. The portfolio return in period $[t+1, t+2]$ after trading costs is then

$$R_{t+2}^P = x_{t+1}^s R_{t+2}^s + x_{t+1}^b r_f - 0.001|x_{t+1}^s - \hat{x}_{t+1}^s| - 0.0002(x_{t+1}^s - 1)^+. \tag{I.5.28}$$

I.6 Empirical Results

I.6.1 Parameter estimation

In this section, the results from the option valuations and portfolio allocations are described. Table B.3 shows the parameter estimates for the five models. The MLE estimates for NAGARCH-P are in-line with the usual results obtained in the literature, with an average GARCH coefficient β of approximately 0.85 and an average ARCH coefficient α of approximately 0.07. The sample mean of the unit risk premium is approximately 0.02. All parameters under MLE are highly significantly[1] different from zero (under the assumed Gaussian distribution with usual significance levels). The NAGARCH-Q estimates echo these results, with all four parameters being of the same magnitude as their equivalents under the physical measure. However, it should be noted that the parameters are no longer as clearly significant as before, as they show much higher standard deviations among the estimates. This is due to the Monte Carlo optimization routine for non-linear functions, which behaves much more volatile than the MLE under the physical measure.

For the unidirectional Markov switching model with two states, the two volatility regimes were estimated together with their transition probabilities (reported are the ones for remaining in the current state). The probabilities are extremely close to 1

[1]Henceforth, '(statistically) significant' is to be understood under the premise of the assumed Gaussian distribution, as well as the usual significance levels.

and show relatively small standard deviations, meaning that the states seem to almost absorb the process after a random amount of time. The bi-directional model with finite states was estimated for $K = 11$ states[2]. The algorithm produces sensible estimates, with β being in-line with the previous estimates and the average volatility estimate over $[t, T]$ matching δ_1 from U-MSM-2, which is a very pleasing result, given that the estimates were obtained using two completely different estimation routines and different start values. A bit surprising are the return-feedback weights q_1 and q_2, which are not significantly different from zero, which would imply that for some cases the uni-directional model may be considered instead, given the latter's much faster and simpler estimation procedure.

Finally, the estimates for the continuous limit model of the B-MSM-K framework, when the number of states approaches infinity, are shown. It is evident that the β estimate is much smaller than in the previous models, which may be due to the return innovations being captured by q_1 and q_2. However, the estimated volatility on day t again matches the previous estimates, although the standard deviation is higher, which is most likely again due to the random-based Monte Carlo method compared to the non-stochastic estimation routines used for U-MSM-2 and B-MSM-11.

I.6.2 Out-of-Sample Option Pricing

After estimating the optimal parameters for each model, the corresponding mean square errors are obtained. Keeping the estimates then constant (as assumed in

[2]It should be noted that although the proposed estimation algorithm is a huge improvement over the initial setting for the purposes of this thesis, running the algorithm to find the best estimates is still very time consuming given the long data period and the large number of $K \times K$ matrices to be stored and multiplied, computational issues with storage and (numerical) loss of significance might occur. Hence, the choice of 11 states as a feasible representative for non-trivial higher state B-MSM.

the respective frameworks), the mean square errors for pricing the call option contracts on the following three Wednesdays are calculated. Table B.4 shows the results[3]. For the finitie-regime Markov Switching models the average mean square error across states is used on each estimation day. As could be expected, the models with more parameters perform better in-sample than the ones with less parameters, which is evident in the smaller average MSE value of B-MSM-Inf compared with NAGARCH-Q (with additionally a smaller standard deviation among these MSE values). The very low B-MSM-11 MSE value, which is lower than the one for infinite states, indicate that the model performs on average better in pricing the options in-sample than the one considering infinite many potential volatility regimes, which shows that a finite-state Markov model is indeed appealing to be considered.
However, out-of-sample the B-MSM-11 model seems to struggle, with higher average mean square errors compared to the other models, suggesting that at least some of its parameters need to be updated over time[4]. The U-MSM-2 model shows a very strong performance out-of-sample, despite its very parsimonious modelling approach, with the evaluation occurring one week later displaying an only slightly higher MSE value compared to NAGARCH-Q and B-MSM-Inf and even a smaller MSE value two weeks out-of-sample. This indicates that the volatility level can be assumed to roughly remain between the two volatility regimes over several weeks, again motivating the class of Markov switching models for option valuation.

[3]It should be noted, that due to the monthly re-estimations of the parameters, which mostly use their previous estimates as start values, and the non-linearity of the optimization routines (generally not guaranteeing a unique optimum), the MSE values can be expected to be higher than if daily re-estimation would have been applied. The trade-off, however, is that the models are fitted based on a longer time period with more market moving events, which is usually desirable.

[4](Duan et al., 2002) update the intercept for each out-of-sample evaluation, which may improve the performance significantly, but loses comparability with U-MSM-2.

I.6.3 Portfolio Strategy Performance

Using (I.3.9), the optimal portfolio weights corresponding to the (physical) risk-premium obtained from the maximum likelihood estimation of NAGARCH-P and the implied volatility estimate induced from each risk-neutral pricing framework are calculated for risk aversion values of 2, 4 and 8. The portfolio is then re-balanced every month using the new estimates, whilst the portfolio value is simulated on a daily basis (with trading costs set as in subsection I.5.3.2). Presented in Table B.5 are the daily portfolio weights held in each strategy. It is very apparent that the weights between NAGARCH-Q, U-MSM-2 and B-MSM-2 only differ ever so slightly, but the weights induced by B-MSM-Inf are noticeably higher on average (due to the lower estimated index volatility, as indicated in Table B.3), but also significantly more volatile. Table B.6 then shows the annualized portfolio returns for each strategy, for the case with trading costs and if trading costs were removed. Given the annualized return of the S&P index over the same period of 16.62%, all models outperform the index in the absence of trading costs for $\kappa = 2$, whilst in their presence it is still achieved by all models besides U-MSM-2, indicating that although the average weights and their standard deviation between U-MSM-2 and NAGARCH-Q are very similar, the strategies differ in their volatility timing. Additionally, all option-pricing models besides U-MSM-2 outperform NAGARCH-P, which demonstrates the value of option-implied information.

Assuming an initial portfolio value of 100, the resulting portfolio performance for each of the strategy and each considered risk-aversion parameter are illustrated in Figure B.4 till Figure B.13, both, when considering trading costs and when omitting trading costs. To measure the Sharpe ratios of the strategies, the average of the short-ends of the LIBOR curve during the data period was taken as a proxy for the daily risk-free rate, which amounts to $3.24E-06$. Table B.7 then shows the Sharpe ratio of the returns of each strategy with and without trading costs. Due to the more volatile returns, B-MSM-Inf no longer performs best under this measure, but instead the NAGARCH models prevail with the highest Sharpe ratios.

Finally, Table B.8 displays the excess annualized certainty equivalent returns (in basis points) for all considered strategies and risk-aversion parameters. Hence, the results can be interpreted as the additional fixed return an investor would accept instead of investing in a portfolio consisting of a long position according to the strategy and a short position in the index. For low risk aversion levels this results in relatively high values due to the outperformances of these strategies over the S&P 500, but for higher risk aversion the negative utility derived from the return variances dominate and the values turn negative.

I.7 Conclusions

This thesis has developed an efficient algorithm for the B-MSM-K models proposed by (Duan et al., 2002) and has inferred model parameters for the bidirectional models with 11 and infinite states and for the unidirectional model with two states from option prices. The results showed strong option pricing performances of the bidirectional MSM, outperforming the NAGARCH model for both 11 and infinite states. Out-of-sample, the performance of the infinite-state B-MSM model was still roughly on par with the NAGARCH-Q model, despite its higher number of parameters that may change between months. The two-state unidirectional model, although not performing as well in-sample, showed promising out-of-sample results, which is noteworthy given the existence of closed-form solutions, making estimation processes comparatively fast. In the second part, the implied volatility estimates were used for portfolio allocation using the two-factor model. Here, both the B-MSM-Inf and B-MSM-11 model showed promising results and the value of option implied volatilities was again shown by comparing the performances to a framework using volatility estimates based on historic return data. It should be noted though that while the average simulated return of the bidirectional models might be appealing for some investors, the returns are also quite volatile, which is shown by their Sharpe ratios and certainty equivalent returns. This may be attributed to the less smooth frameworks when compared to the NAGARCH models.

Based on the promising results for option pricing and asset allocation, it is up to future research to verify these results for different markets and different numbers of states considered. Additionally, the CBOE Volatility Index (VIX) could be incorporated as a secondary measure for implied volatility and used as a control variable in the various estimation procedures, as it obtains the expected volatility of the S&P 500 over the next 30 days and thus makes an logical inclusion.

Part II

Practical Alternatives: (Ad-hoc) Black-Scholes and the VIX

II.1 Motivation

Despite the empirical success of the models in Part I, namely NAGARCH and the bidirectional Markov switching models of (Duan et al., 2002), they do have some practical drawbacks. The relatively complex, computationally burdensome and time intensive estimation routines, such as (randomized quasi) Monte Carlo methods and multinomial lattices, discourage some financial institutions and (private) investors from using these models. These practitioners often depend on an instantaneously available estimate and cannot afford to spend time on estimation routines. Therefore, models that produce closed-form solutions for pricing securities such as derivatives are in high demand.

One of the most well-known of these models is the Nobel prize-winning Black-Scholes model (Black and Scholes, 1973; Merton, 1973). However, in the past decades, several studies in academia have shown that the constant volatility assumption of Black-Scholes does not hold up empirically and have thus looked for model extensions and alternative frameworks. One of these model extensions, known as the ad-hoc Black-Scholes model (Dumas et al., 1998), has shown very promising results and was therefore adopted by many practitioners and researchers alike.

Despite pricing derivatives, volatility is also used in many other ways, including risk and portfolio management. Hence, this thesis adopts the modern portfolio theory of (Markowitz, 1952; Markowitz, 1959) and uses the implied volatility under the ad-hoc Black-Scholes model as an input parameter for the risk component of the framework. Two, alternative strategies are considered as benchmarks in the portfolio analysis part; a buy-and-hold strategy and one using the VIX. The CBOE Volatility Index (VIX) also estimates implied volatility and provides investors with an easily and instantaneously available estimate.

The second thesis part is structured in the following way. First, in chapter II.2, the portfolio framework, the ad-hoc Black-Scholes model, and the VIX are presented. Then, in chapter II.3 the data set used for the empirical analysis is described, before, in chapter II.4, the analysis settings and results are reported[1]. The thesis ends in chapter II.5 with the conclusions of the empirical analysis and gives an outlook for potential further research.

[1]Since the second part is required to exist independently of the first part, the reader should be aware that some of the presented content, such as the data description and the empirical settings, are summaries of the corresponding chapters in Part I.

II.2 Methodology

II.2.1 Portfolio Theory

As in Part I, modern portfolio theory (MPT) is employed. MPT is based on the Nobel prize-winning two-factor model proposed in (Markowitz, 1952) and linked to the utility theory of (Von Neumann and Morgenstern, 1945) by (Markowitz, 1959). Tobin's separation theorem (Tobin, 1958) then states that the investment and financing decision of the investor are independent of each other. Hence, the optimal portfolio in this framework is a split between the *market portfolio* and a *risk-free* asset, and only the amounts invested in each of the two assets depend on the risk aversion level of the investor.

Hence, it is assumed that an investor can invest in two assets, a risky one, namely the S&P 500[1], and a risk-free one such as a 10-year U.S. government bond[2]. The investor is assumed to be equipped with an information set Ψ_t that contains all relevant information until time t. The challenge for the investor is then to allocate the fractions of his or her portfolio to the risk-free and risky assets in a manner that maximizes the investor's utility. To simplify, it is assumed that the investor's utility function only relies on the mean and the variance of the portfolio returns.

[1]Since buying the whole index is not very practically, the usual way this theory can nowadays be implemented is by buying an Exchange Traded Fund (ETF) on the stock index. However, this might lead to tracking errors and the portfolio performance might deviate even in the case when all assumptions were met.

[2]Although the 10-year U.S. government bond might not be entirely free of risk, especially during the current period of increased political risk, it is still frequently used in academia as a synonym for a risk-free investment.

Let the fractions of the portfolio invested in each asset be denoted by x^s and x^b, respectively. The expected value and variance of the portfolio returns are given by[3]:

$$\begin{aligned}&\mathrm{E}_t[R^P_{t+1}] = x^s_t \mathrm{E}_t[R^s_{t+1}] + x^b_t r_f \text{ and}\\&\mathrm{V}_t[R^P_{t+1}] = (x^s_t)^2 \mathrm{V}_t[R^s_{t+1}],\end{aligned} \tag{II.2.1}$$

with the usual short notation $\mathrm{E}_t[R^P_{t+1}] \equiv \mathrm{E}[R_{t+1}|\Psi_t]$ and r_f being the appropriate risk-free rate, i.e. the current risk-free yield corresponding to the time horizon of the investment.

The investor's optimal allocation at time t under this framework thus needs to simultaneously maximize $\mathrm{E}_t[R^P_{t+1}]$ and minimize $\mathrm{V}_t[R^P_{t+1}]$. This is formalized as

$$x^*_t = arg \min_{x^s_t, x^b_t} -\kappa \left(x^s_t \mathrm{E}_t[R^s_{t+1}] + x^b_t r_f \right) + \frac{1}{2}(x^s_t)^2 \mathrm{Var}_t[R^s_{t+1}] \quad \text{s.t.} \quad x^s_t + x^b_t = 1. \tag{II.2.2}$$

with κ and $x^s_t + x^b_t = 1$ being the investor's risk aversion level and budget constraint, respectively.

As can be shown, this leads to the *parametric-efficient* portfolio allocations determined by

$$x^{s*}_t = \frac{\mathrm{E}_t[R^s_{t+1}] - r_f}{\kappa \mathrm{Var}_t[R^s_{t+1}]}, \quad x^{b*}_t = 1 - x^{s*}_t. \tag{II.2.3}$$

Thus, estimates for the mean and variance of the return process are needed.

II.2.2 Black-Scholes

Although the mean of a return series has usually been impossible to forecast accurately, researchers have been successful in predicting the second moment to a certain degree. Throughout the years, a great number of papers attempted to formalize increasingly convoluted frameworks in an attempt to fit cross-sections of observed

[3]For the derivation of these results, see for example (Best, 2010).

option prices as perfectly as possible upon a proposed volatility process[4]. However, quite often some of these approaches face serious over-fitting issues and thus usually do not perform well out-of-sample. One very successful and surprisingly simple model that is regularly being used by some market participants is the ad-hoc Black-Scholes model (ABS).

In order to present ABS, it is necessary to give a short overview of the original Nobel prize-winning framework presented in the seminal papers (Black and Scholes, 1973) and (Merton, 1973). Similarly to the setting assumed in this thesis, the Black-Scholes model (henceforth, BS) considers the availability of two securities, a risk-free investment B, say a government bond, and an investment opportunity S with a volatile return history, say a stock (index). Due to its risk-free characteristic, the price of B at time t is given by

$$B(t) = e^{r_f t}. \tag{II.2.4}$$

Meanwhile, S (under the historical probability measure $\mathbb{P}$) is assumed to follow a geometric Brownian motion, i.e.

$$\frac{dS(t)}{S(t)} = \mu dt + \sigma dW^{\mathbb{P}}(t), \tag{II.2.5}$$

where $dW^{\mathbb{P}}$ is a standard Brownian motion under $\mathbb{P}$ with respect to a given filtration induced by the information set of the investor. The parameters μ and σ are usually referred to as *drift* and *volatility* of the process and resemble the mean of the instantaneous rate of return on the stock and its sensitivity to the risk factor $dW^{\mathbb{P}}$.

To solve the differential equation (II.2.5), Itô's lemma (Itô, 1944) is applied, which leads to

$$S_t = S_0 e^{(\mu - \frac{1}{2}\sigma^2)t + \sigma W_t^{\mathbb{P}}}, \tag{II.2.6}$$

[4]Most popular are probably the stochastic volatility models, such as (Heston and Nandi, 2000), and the deterministic volatility models, which are usually built around the GARCH framework (Engle, 1982; Bollerslev, 1986). For a brief overview and discussion of the main modeling approaches, see section I.2.2 of Part I.

where S_0 denotes the start value of the stock price process.

To price a derivative under this framework, it is necessary to derive the price process under an equivalent risk-neutral probability measure $\mathbb{Q}$. It can be shown that this yields

$$S_t = S_0 e^{(r_f - \frac{1}{2}\sigma^2)t + \sigma W_t^{\mathbb{Q}}}, \tag{II.2.7}$$

with r_f denoting the risk-free rate and $W_t^{\mathbb{Q}} = W_t^{\mathbb{P}} + \frac{\mu - r_f}{\sigma} t$ being a standard Brownian motion under $\mathbb{Q}$. Black, Scholes and Merton proved that pricing a European call option on a non-dividend paying stock and a European put option on a non-dividend paying stock, with maturity T and strike price K, has the following closed-form solutions:

$$C_t = S_t \Phi(d_1) - K e^{-r_f (T-t)} \Phi(d_2) \tag{II.2.8}$$

and

$$P_t = K e^{-r_f (T-t)} \Phi(-d_2) - S_t \Phi(-d_1), \tag{II.2.9}$$

where

$$\begin{aligned} d_1 &= \frac{\ln(S_t/K) + (r_f + \sigma^2/2)(T-t)}{\sigma\sqrt{T-t}}, \\ d_2 &= d_1 - \sigma\sqrt{T-t}. \end{aligned} \tag{II.2.10}$$

As usual, $\Phi(x)$ is the distribution function of a standard normal random variable, i.e.

$$\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy. \tag{II.2.11}$$

The Black-Scholes option price can thus be expressed as $BS(S_t, K, T, r_f, \sigma)$, i.e. as a function of the stock price at time t, the option's strike price and maturity, the risk-free rate and the stock's volatility. Two important observations must be made. First, since the first four parameters are given at time t, the option pricing under Black-Scholes boils down to estimating the unobservable stock volatility. Second, under

BS, the stock volatility is assumed to be constant over time and to be the same for all option contracts.

This leads to the following pricing procedure under BS:

1. Obtain a cross-section of option prices on day t, C^{mkt}.
2. For each option find the implied volatility $\sigma_{BSIV,i}$, i.e. $\sigma_{BSIV,i} : C^{mkt}(K_i, T_i) = BS(S_t, K, T, r_f, \sigma_{BSIV,i})$.
3. The Black-Scholes implied volatility on day t is then estimated as $\sigma_{BSIV} = \frac{1}{N} \sum_{i=1}^{N} \sigma_{BSIV,i}$.
4. Hence, the BS price of option i on day t is $BS(S_t, K_i, T_i, r_f, \sigma_{BSIV})$ and the BS price of option j on day $t+1$ (out-of-sample) is $BS(S_{t+1}, K_j, T_j, r_f, \sigma_{BSIV})$.

II.2.3 Ad-hoc Black-Scholes

Nowadays, BS is widely rejected by academia and many practitioners, because the constant volatility assumption is considered to be unfit with many market observations[5]. A model extension that deals with this issue was presented in (Dumas et al., 1998), where the volatility is still considered to be deterministic and the BS pricing formulas are applied, but each option uses its own volatility rather than keeping the same value over all different strike prices and maturities. Under this ad-hoc Black-Scholes model (henceforth, ABS), the BS implied volatilities are regressed upon a polynomial function of the strike prices and maturities of these options. Although there are infinite possible polynomial functions that could be used for this matter, (Dumas et al., 1998) suggest a quadratic form based on the observed parabolic Black-Scholes implied volatility patterns ("volatility smile/skew"). Hence, ABS assumes

[5] According to Black himself a few years later: "... if the volatility of a stock changes over time, the option formulas that assume a constant volatility are wrong." (Black, 1976, p. 177).

the following structure for the volatility of option i with strike price K_i and maturity T_i:

$$\sigma_{BSIV,i} = \beta_0 + \beta_1 K_i + \beta_2 K_i^2 + \beta_3 T_i + \beta_4 T_i^2 + \beta_5 K_i T_i + \varepsilon_i, \tag{II.2.12}$$

where ε_i is a zero mean error term. Note that $\beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5$ leads to the original Black-Scholes model.

The pricing procedure under ABS can thus be summarized as:

1. Obtain a cross-section of option prices on day t, C^{mkt}.
2. For each option find the implied volatility $\sigma_{BSIV,i}$, i.e. $\sigma_{BSIV,i} : C^{mkt}(K_i, T_i) = BS(S_t, K, T, r_f, \sigma_{BSIV,i})$.
3. Run the regression in (II.2.12) and denote the fitted values by $\hat{\sigma}_{ABS}$.
4. The ABS price of option i on day t is then $BS\left(S_t, K_i, T_i, r_f, \hat{\sigma}_{ABS} K_i, T_i)\right)$ and for option j on day $t+1$ (out-of-sample): $BS\left(S_{t+1}, K_j, T_j, r_f, \hat{\sigma}_{ABS}(K_j, T_j)\right)$.

II.2.4 The VIX

An alternative way to measure the S&P 500 stock index volatility is the Volatility Index (VIX) published by the Chicago Board Options Exchange (CBOE). The VIX, often also called the 'investor fear gauge' (Whaley, 2000), provides a measure for the expected volatility of the S&P 500 over the next 30 days and is frequently recognized as one of the chief benchmarks for stock market volatility in the U.S. Its nickname reflects the index' tendency to jump up when a financial crisis is imminent, as can be observed in Figure B.18. The observation is further supported by a correlation analysis, where the VIX showed a correlation of -0.6851 with the S&P 500 index during the period January 01, 2010 to July 31, 2014, providing further evidence for the VIX's importance for S&P 500 investors.

Similar to the BS and ABS models above, the VIX provides an implied volatility estimate, which is inferred from cross-sections of observed option prices. According to (Chicago Board Options Exchange, 2014), the VIX is calculated as

$$\hat{\sigma}_{VIX,t} = 100\sqrt{\frac{2}{T}\sum_i \frac{\Delta K_i}{K_i^2} e^{r_f T} Q_t(K_i) - \frac{1}{T}\left[\frac{F_t}{K_0}\right]^2}, \tag{II.2.13}$$

where K and T again denote the strike price and time to maturity of an option and r_f the appropriate risk-free rate. Additionally, F stands for the forward value of the S&P 500 index, K_0 for the first strike price below F and $Q(K_i)$ for the midpoint of the bid-ask spread for each option with strike K_i. To interpolate the 30-day horizon, the VIX uses put and call options with more than 23 days and less than 37 days to expiration[6]. One major advantage of the VIX is the availability of historical data for the last three decades provided by CBOE. Although the VIX does not attempt to be an instrument to price option contracts (like the BS and ABS models do), because of its simplicity and high data availability, it is a premier source for obtaining estimates for the volatility of the S&P 500 and hence suitable for the portfolio allocation analysis part of this thesis.

[6]For further details on the calculation of the VIX, see (Chicago Board Options Exchange, 2014).

II.3 Data

The thesis uses weekly stock market data on the S&P 500 index obtained from the Wharton Research Data Services of the University of Pennsylvania. The data set was obtained for the period January 01, 2010 to July 31, 2014[1], and on each day, the closing index level is obtained (Figure B.1). Furthermore, data on all traded put and call option contracts (with the index as the underlying) are collected every Wednesday[2], leading to 239 observation days. The option data set consists of the best bid and ask prices, trading volumes, strike prices, and maturities of the option contracts.

Since the S&P 500 index options are European style and many of the index's stocks pay dividends, adjustments need to be made. Therefore, the cash dividend payments are obtained from the S&P 500 information bulletin and the linearly interpolated LIBOR curve is used as a proxy for the risk-free rate. Given both data sets, and assuming the future dividend payment structure to be known at time t, the present value of all future dividend payments until the maturity of each option is calculated and subsequently subtracted from the current stock value. Note that this leads to an adjusted index level on each day that might differ within each cross-section of option contracts.

The option data set is then filtered in several ways (following, in part, (Dumas et al., 1998)). Options with maturities of less than six or more than 100 days are excluded,

[1] For detailed reasoning on why the S&P 500 index and this particular period length were chosen, as well as many further details, it is referred to chapter I.4 of Part I.

[2] Wednesday was chosen, as usual, because it historically has been the weekday with the least holidays in a year in the U.S. In any case, if a holiday still falls on a Wednesday, the next available trading day is used instead.

as well as options with a trading volume of less than 50, to avoid liquidity-related biases in the observed option prices. Additionally, options with prices below $\$0.30$ are filtered out to prevent problems stemming from price discreteness. Next, a moneyness filter is applied to remove deep-in-the-money and deep-out-of-the-money options. Finally, a set of non-arbitrage conditions (as discussed in (Gonçalves and Guidolin, 2006)) is tested and violations extracted from the data series (for further details see section I.4.3 of Part I). To make the results comparable to other well-known papers on option pricing (such as (Christoffersen and Jacobs, 2004)), only call option contracts are considered, although the analysis can be performed analogously with put option contracts.

II.4 Empirical Analysis

The weekly recordings of option prices are split into four sections. The first week's observations serve for the in-sample parameter estimation and the following three weeks are used for out-of-sample testing. Option contracts are then classified by their maturity and moneyness, with the numbers of contracts and their prices displayed in Table B.1 and Table B.2, respectively.

II.4.1 Parameter Estimation

Following the pricing procedures outlined in chapter II.2 on a cross-section of option prices, the parameters of the BS and ABS models are estimated by minimizing the mean square error with the observed market prices, i.e. for ABS:

$$MSE_t = \frac{1}{N} \sum_{j=1}^{N} \left(C_{j,t}^{ABS} - C_{j,t}^{mkt} \right)^2 . \tag{II.4.1}$$

These parameters are then held constant over the next four weeks and the mean square errors of the modeled option prices in each of the three out-of-sample weeks are calculated.

For the portfolio allocation analysis, and building upon the results in Part I, the unit risk premium during the period January 01, 2010 until July 31, 2014, is estimated at 2.41% (see λ in Table B.3). As an estimate for the daily volatility of the S&P 500 over the next period, the mean of $\hat{\sigma}_{ABS}$ is used, which is obtained during each in-sample analysis every four weeks. For the risk aversion level, the thesis considers three cases, namely $\kappa = 2, 4, 8$. The portfolio framework then calculates the parametric-efficient weights via (II.2.3) and re-balances monthly when the new weights are set[1].

II.4.2 Trading Costs

Other than in the asset allocation model outlined in section II.2.1, the thesis additionally considers the effect of trading costs on the portfolio performance[2]. Furthermore, short-selling is allowed at the risk-free rate but capped at 200% of equity, resulting in $x^s \in [0, 2]$ and $x^b \in [-1, 1]$. Short selling costs apply and are set at two basis points of borrowed value, but transaction costs for trading the risk-free rate are omitted. For the risky asset, (Domowitz et al., 2001) estimate the combined transaction costs for trading on the NYSE in 2000 at roughly 30 basis points of trade value. Thus, considering significant innovations in infrastructure and technology during the last two decades, transaction costs are set at ten basis points of trade value.

To determine the traded portfolio value, it is necessary to determine by how much the previously allocated fraction of the portfolio has changed. Hence, at time t the current fraction of the portfolio invested in the risky asset is

$$\hat{x}_t^s = \frac{x_{t-1}^s e^{R_t^s}}{x_{t-1}^s e^{R_t^s} + x_{t-1}^b e^{r_f}}, \tag{II.4.2}$$

where R_t^s and r_f are the returns of the risky and risk-free assets over $[t-1, t]$, respectively. The portfolio return over period $[t, t+1]$ after trading costs is then given

[1]Although daily re-balancing is quite common in academia, as it more accurately displays the performance of the strategy, it is very uncommon in the markets due to several market frictions making it quite expensive.

[2]For a more in-depth discussion of the proposed trading costs, see subsection I.5.3.2 of Part I.

by[3]

$$R_{t+1}^{P} = x_t^s R_{t+1}^s + x_t^b r_f - 0.001|x_t^s - \hat{x}_t^s| - 0.0002\left(x_t^s - 1\right)^+ . \tag{II.4.3}$$

II.4.3 Performance Measures

Besides obtaining the mean and variance of the portfolio returns, performance is also measured by the *Sharpe Ratio* (SR) and the *Certainty Equivalent Return* (CER). Whilst SR measures the return the portfolio strategy provides in excess of the risk-free alternative in relation to the risk taken, it does not incorporate the investor's level of risk aversion. This is included by adding CER, which gives an estimate for the yield the investor would be willing to accept for opting not to invest in the portfolio strategy.

Denoting the mean and variance of the resulting portfolio returns by $\hat{\mu}$ and $\hat{\sigma}$, the two measures are calculated as

$$SR_j = \frac{\hat{\mu}_j - r_f}{\hat{\sigma}_j} \tag{II.4.4}$$

and

$$CER_j = \hat{\mu}_j - \frac{\kappa}{2}\hat{\sigma}_j{}^2. \tag{II.4.5}$$

II.4.4 Results

The ad-hoc Black-Scholes model was introduced and is being used as an improvement over the standard Black-Scholes model. Hence, the analysis described so far

[3]By usual definition, $\left(x_t^s - 1\right)^+ \equiv \max\left(x_t^s - 1, 0\right)$.

is compared to the BS framework, the VIX framework as another practical implementation, and to the models described in Part I, namely NAGARCH and the bidirectional Markov switching models of (Duan et al., 2002).

The parameter estimates for the BS and ABS models are presented in Table B.12. It is very apparent that none of the ABS estimates are significantly (at common significance levels) different from zero, while the BS estimate is significantly different from zero. However, this is of relatively low concern for the goals of this paper as it focuses on option pricing rather than factor analysis. In fact, the in-sample and out-of-sample pricing results for ABS, presented in Table B.9, show much smaller mean square errors than the Black-Scholes model. Due to its approach, ABS achieves very strong in-sample pricing accuracy, despite its simple and very fast estimation procedure. Furthermore, as displayed in Figure B.19, the ABS manages to replicate a *volatility skew*, the phenomenon (frequently observed for index options) where implied volatility is higher for options with lower strike prices than for options with higher strike prices (at the same maturity)[4]. However, out-of-sample, although still clearly surpassing standard Black-Scholes, the model performs worse than the NAGARCH and Markov switching models of Part I, as it produces higher out-of-sample mean square errors.

For the second part of the analysis, the portfolio strategy, the ABS induced portfolios are compared to the VIX induced portfolios, according to the equations developed under modern portfolio theory in chapter II.2. This attempts to show that the extra effort that ABS requires (compared to BS or even more when compared to VIX) can be beneficial even from an asset management perspective. The results are quantified in Table B.10 and visualized in Figures B.14-B.17.

[4] A popular explanation for this kind of observation is that in-the-money calls (i.e. calls with lower strike prices) are being bought to leverage stock exposure instead of buying the index itself, in order to increase return on investment.

The portfolio weights under the ABS strategy are quite aggressive, in the sense that for $\kappa = 2$ the portfolio holds an average weight of approximately 144% of the index, i.e. is constantly leveraging its exposure. For this low risk-aversion level, the ABS portfolio strategy then achieves a 2% to 6% (with/without trading costs) higher annualized return than a straight buy-and-hold strategy of the index would have achieved over the same period. However, these returns appear to also be a bit more volatile, so that only without trading costs does the strategy achieve a Sharpe ratio and Certainty Equivalent Return (0.8072 and 1.46%, respectively) higher than that of a buy-and-hold strategy (0.7069 and 1.13%, respectively). The VIX strategy instead uses a less aggressive strategy, with smaller fractions of the portfolio allocated to the stock index and a higher fraction invested in the risk-free asset. This indicates, based on (II.2.3), that the VIX implied volatility estimate is higher than its ABS counterpart, an observation that has also been made when comparing the VIX to the GARCH framework (Hao and Zhang, 2013). This results in lower annualized returns, however, the annualized return of the strategy for low risk aversion ($\kappa = 2$) is still higher than the one of a buy-and-hold strategy. As with ABS, VIX also does manage to achieve a higher SR (0.8066) and CER (1.24%) in this case, when compared to the buy-and-hold strategy. However compared with the ABS model the VIX strategy performs worse in almost all measurements, with almost all of the Sharpe ratios and Certainty Equivalent Returns being lower than in the ABS case.

If one compares these results with the ones obtained in Part I, it becomes apparent that the NAGARCH model, which has seen remarkable popularity and empirical success during recent years in many academic papers, still performs noticeably better than even the ABS model. Specifically, the NAGARCH(-Q)'s Sharpe ratios for the well-performing case of $\kappa = 2$ are 0.8387 and 0.9381 (with/without trading costs), with a similar picture drawn by the Certainty Equivalent Returns. It is worth noting that according to the SR and CER measures, NAGARCH performs better than ABS (and therefore VIX) also for higher risk aversion levels (in fact, NAGARCH manages to achieve a higher Sharpe ratio and Certainty Equivalent Return than a buy-and-hold strategy even for $\kappa = 4$, which is not the case with ABS).

Generally, the more risk averse the investor is, the less is being invested in the risky asset, which during the analyzed period that showed a clear uptrend of the S&P 500, hurts the strategy's performance but might protect it during times of crises. Because of this, the strategies corresponding to higher risk aversion levels of both portfolio strategy can no longer compete with a buy-and-hold investor during the studied period. Thus, based on these results, the more risk averse an investor is, the less he or she should try to time market volatility with this strategy (especially during times of prosperity) and rather straight up hold the index, which is an intuitive result and frequently replicated in studies on portfolio management.

II.5 Conclusions

After Part I has analyzed the economic value of implied volatilities in a modern portfolio theory (MPT) setting, by comparing the NAGARCH framework with the bidirectional Markov switching models proposed in (Duan et al., 2002), Part II has built the bridge to models and instruments frequently favored by practitioners. Since the results of MPT require an estimate for the risky asset's volatility (in this case of the stock index), a major part of this thesis deals with option pricing to determine which model produces the lowest pricing errors, and thus potentially the best volatility estimate.

Starting from the Nobel prize-winning Black-Scholes (BS) model, which, in the past, had been incredibly popular with practitioners, the ad-hoc Black-Scholes (ABS) model of (Dumas et al., 1998) was introduced as a model extension and compared to its predecessor in form of option pricing. The empirical analysis confirms the benefit of the ABS model as it achieves smaller mean square pricing errors than the BS model and replicates the *volatility smirk* that has been frequently observed with index options.

Then, based on the ABS implied volatility estimates a portfolio is constructed under MPT and its performance compared with a buy-and-hold strategy. Despite showing a clear uptrend, the ABS strategy achieves a higher Sharpe ratio and Certainty Equivalent Return for low risk aversion, highlighting the value of volatility timing for asset management purposes. Additionally, the potential value of the increased complexity of ABS was shown again by comparing the model to a strategy based on the popular CBOE Volatility Index (VIX), which does not require any further computations and is readily available, but seems to overestimate the index volatility relative to the ABS (and GARCH) estimates. The results show that an investor who wants to use either of the two instruments for a quantitative portfolio strategy using MPT would likely be better off with the ad-hoc Black-Scholes model.

As was shown, if the investor is willing to put even more effort into his or her quantitative modeling approach, there have been significant contributions in recent years in academia that might be able to achieve even better results. Most notably, the NA-GARCH framework seems to still perform considerably better both for option pricing and during the portfolio analysis under MPT, when compared to ABS. However, this comes at the cost of increased complexity and model risk, as was shown during the discussion of the randomized quasi Monte Carlo simulation for parameter estimation in Part I. Therefore, using the ad-hoc Black-Scholes model might be a sensible approach for practitioners (e.g. private investors) who want to improve upon (or get an alternative for) the VIX and/or the standard Black-Scholes model and do not want to indulge in more theoretical, and potentially complicated, frameworks.

Regarding an outlook for future research, it is to be shown that these results can hold up during a different period, most importantly during a downward cycle, and perhaps in a different market. It should also be noted that this analysis assumes the appropriateness of the MPT formulas. Despite its continuing relevance in academia, some alternatives to MPT have become popular in recent years, such as the post-modern portfolio theory, which softens some of the assumptions of MPT (such as the assumed symmetry of the return distribution) or the Black-Litterman model, which uses less assumptions and incorporates the individual investor's belief in his or her views. Hence, it is to be shown that the results in this thesis also hold up under an alternative portfolio theory.

References

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

Bams, D., Lehnert, T., and Wolff, C., 2004. “Loss Functions in Option Valuation: A Framework for Model Selection”, *CEPR Discussion Paper*.

Best, M., 2010. *Portfolio optimization*. Chapman & Hall/CRC finance series. London: Chapman & Hall.

Black, F., 1976. “Studies of Stock Price Volatility Changes”, *Proceedings of the Business and Economics Section of the American Statistical Association*, 177–181.

Black, F. and Scholes, M., 1973. “The Pricing of Options and Corporate Liabilities”, *Journal of Political Economy*, vol. 81, 637–654.

Bollen, N., 1998. “Valuing Options in Regime-Switching Models”, *The Journal of Derivatives*, vol. 6, 38–49.

Bollen, N., Gray, S., and Whaley, R., 2000. “Regime-Switching In Foreign Exchange Rates: Evidence From Currency Option Prices”, *Journal of Econometrics*, vol. 94, 239–276.

Bollerslev, T., 1986. “Generalized autoregressive conditional heteroskedasticity”, *Journal of Econometrics*, vol. 31, 307–327.

Busch, T., Christensen, B., and Nielsen, M., 2011. “The Role of Implied Volatility in Forecasting Future Realized Volatility and Jumps in Foreign Exchange, Stock, and Bond Markets”, *Journal of Econometrics*, vol. 160, 48–57.

Cai, J., 1994. "A Markov Model of Switching-Regime ARCH", *Journal of Business & Economic Statistics*, vol. 12, 309–316.

Chicago Board Options Exchange, 2014. *The CBOE Volatility Index - VIX: White Paper*.

Christoffersen, P. and Jacobs, K., 2002. "Which Volatility Model for Option Valuation?", *EFA 2002 Berlin Meetings Discussion Paper*.

Christoffersen, P. and Jacobs, K., 2003. "The Importance of the Loss Function in Option Valuation", *EFA Annual Conference Paper*.

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

Cox, J., Ross, S., and Rubinstein, M., 1979. "Option pricing: A simplified approach", *Journal of Financial Economics*, vol. 7, 229–263.

Domowitz, I., Glen, J., and Madhavan, A., 2001. "Global equity trading costs", *ITG White Paper*, vol. May 8.

Duan, J.-C., 1995. "The Garch Option Pricing Model", *Mathematical Finance*, vol. 5, 13–32.

Duan, J.-C., Popova, I., and Ritchken, P., 2002. "Option Pricing under Regime Switching", *Quantitative Finance*, vol. 2, 116–132.

Duan, J.-C. and Simonato, J.-G., 1998. "Empirical Martingale Simulation for Asset Prices", *Management Science*, vol. 44, 1218–1233.

Dueker, M., 1997. "Markov Switching in GARCH Processes and Mean-Reverting Stock-Market Volatility", *Journal of Business & Economic Statistics*, vol. 15, 26.

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

Dupire, B., 1994. "Pricing with a smile", *Risk*, vol. 7, 18–20.

Engle, R., 1982. "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation", *Econometrica*, vol. 50, 987.

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

Faure, H. and Tezuka, S., 2002. "Another Random Scrambling of Digital (t,s)-Sequences", *Monte Carlo and quasi-Monte Carlo methods 2000*. Ed. by K.-T. Fang, F. Hickernell, and H. Niederreiter. Berlin and New York: Springer, pp. 242–256.

Fleming, J., Kirby, C., and Ostdiek, B., 2001. "The Economic Value of Volatility Timing", *The Journal of Finance*, vol. 56, 329–352.

Giudice, G., 2017. "Assessing the Economic Value of Option-Implied GARCH Estimates: A Portfolio Approach". Master's thesis. Milan, Italy: Bocconi University.

Glosten, L., Jagannathan, R., and Runkle, D., 1993. "On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks", *The Journal of Finance*, vol. 48, 1779–1801.

Gonçalves, S. and Guidolin, M., 2006. "Predictable Dynamics in the S&P 500 Index Options Implied Volatility Surface", *The Journal of Derivatives*, vol. 79, 1591–1635.

Gray, S., 1996. "Modeling the conditional distribution of interest rates as a regime-switching process", *Journal of Financial Economics*, vol. 42, 27–62.

Guidolin, M., 2011. "Markov Switching Models in Empirical Finance", *Missing data methods: Time-series methods and applications*. Ed. by D. Drukker. Advances in econometrics. Bingley: Emerald Group Publishing Limited, pp. 1–86.

Guidolin, M., 2013. "Markov Switching Models in Asset Pricing Research", *Handbook of Research Methods and Applications in Empirical Finance*. Ed. by A. Bell, C. Brooks, and M. Prokopczuk. Handbooks of research methods and applications. Edward Elgar, pp. 3–44.

Guidolin, M. and Timmermann, A., 2007. "Asset allocation under multivariate regime switching", *Journal of Economic Dynamics and Control*, vol. 31, 3503–3544.

Häfner, R., 2004. *Stochastic implied volatility: A factor-based model / Reinhold Hafner*. Vol. 545. Lecture notes in economics and mathematical systems, 0075-8442. Berlin: Springer.

Hamilton, J., 1988. "Rational-expectations econometric analysis of changes in regime", *Journal of Economic Dynamics and Control*, vol. 12, 385–423.

Hamilton, J., 1989. "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle", *Econometrica*, vol. 57, 357.

Hamilton, J., 2005. *What's Real About the Business Cycle?* Cambridge, MA: National Bureau of Economic Research.

Hamilton, J. and Susmel, R., 1994. "Autoregressive Conditional Heteroskedasticity and Changes in Regime", *Journal of Econometrics*, vol. 64, 307–333.

Hao, J. and Zhang, J., 2013. "GARCH Option Pricing Models, the CBOE VIX, and Variance Risk Premium", *Journal of Financial Econometrics*, vol. 11, 556–580.

Heston, S. and Nandi, S., 2000. "A Closed-Form GARCH Option Valuation Model", *Review of Financial Studies*, vol. 13, 585–625.

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

Itô, K., 1944. "Stochastic integral", *Proceedings of the Imperial Academy*, vol. 20, 519–524.

Jeanne, O. and Masson, P., 2000. "Currency crises, sunspots and Markov-switching regimes", *Journal of International Economics*, vol. 50, 327–350.

Keim, D. B. and Madhavan, A., 1998. "The Cost of Institutional Equity Trades", *Financial Analysts Journal*, vol. 54, 50–69.

Klaassen, F., 2002. "Improving GARCH volatility forecasts with regime-switching GARCH", *Advances in markov-switching models*. Ed. by J. Hamilton and B. Raj. Studies in empirical economics. Heidelberg: Physica-Verlag, pp. 223–254.

Kolusheva, D., 2008. "Out-of-Sample Performance of Asset Allocation Strategies", *SSRN Electronic Journal*.

Kostakis, A., Panigirtzoglou, N., and Skiadopoulos, G., 2011. "Market Timing with Option-Implied Distributions: A Forward-Looking Approach", *Management Science*, vol. 57, 1231–1249.

Lehar, A., Scheicher, M., and Schittenkopf, C., 2002. "GARCH vs. stochastic volatility: Option pricing and risk management", *Journal of Banking & Finance*, vol. 26, 323–345.

Mandelbrot, B., 1963. "The Variation of Certain Speculative Prices", *The Journal of Business*, vol. 36, 394–419.

Markowitz, H., 1952. "Portfolio Selection.", *The Journal of Finance*, vol. 7, 77–91.

Markowitz, H., 1959. *Portfolio selection: Efficient diversification of investments*. Vol. 16. Cowles Foundation monograph.

Marquering, W. and Verbeek, M., 1999. "An empirical analysis of intertemporal asset pricing models with transaction costs and habit persistence", *Journal of Empirical Finance*, vol. 6, 243–265.

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

Merton, R., 1976. "Option pricing when underlying stock returns are discontinuous", *Journal of Financial Economics*, vol. 3, 125–144.

Mitra, S., 2011. "A Review of Volatility and Option Pricing", *International Journal of Financial Markets and Derivatives*, vol. 2, 149–179.

Natenberg, S., 1994. *Option Volatility & Pricing: Advanced Trading Strategies and Techniques*. McGraw Hill Professional.

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

Owen, A., 1998. “Scrambling Sobol’ and Niederreiter–Xing Points”, *Journal of Complexity*, vol. 14, 466–489.

Renault, E., 1997. “Econometric models of option pricing errors”, *Econometric Society Monographs*, vol. 28, 223–278.

Renault, E. and Touzi, N., 1996. “Option Hedging And Implied Volatilities In A Stochastic Volatility Model”, *Mathematical Finance*, vol. 6, 279–302.

Ritchken, P. and Trevor, R., 1999. “Pricing Options under Generalized GARCH and Stochastic Volatility Processes”, *The Journal of Finance*, vol. 54, 377–402.

Rubinstein, M., 1994. “Implied Binomial Trees”, *The Journal of Finance*, vol. 49, 771–818.

Sims, C. and Zha, T., 2004. “Were there Regime Switches in U.S. Monetary Policy?”, *FRB of Atlanta Working Paper*, vol. 2004.

Sobol, I., 1967. “On the distribution of points in a cube and the approximate evaluation of integrals”, *UDSSR Computational Mathematics and Mathematical Physics*, vol. 7, 86–112.

Stoll, H., 1995. “The Importance of Equity Trading Costs: Evidence from Securities Firms’ Revenues”, *Global equity markets: Technological, competitive, and regulatory challenges*. Ed. by R. Schwartz. New York: New York University Salomon Center.

Tobin, J., 1958. “Liquidity Preference as Behavior Towards Risk”, *The Review of Economic Studies*, vol. 25, 65–86.

Von Neumann, J. and Morgenstern, O., 1945. *Theory of Games and Economic Behavior*. Princeton, NJ: Princeton University Press.

Whaley, R., 2000. “The Investor Fear Gauge”, *The Journal of Portfolio Management*, vol. 26, 12–17.

Zakoian, J.-M., 1994. “Threshold heteroskedastic models”, *Journal of Economic Dynamics and Control*, vol. 18, 931–955.

A Mathematical Appendix

A.1 Derivation of B-MSM-Inf

Proof. Given the B-MSM formalization under the physical measure $\mathbb{P}$ in (I.3.17), the volatility process for $t = 1, \ldots, T-1$ and $1 < i < K$ can equivalently be written as

$$\begin{aligned}
&\sigma_{t+1}^{(K)} = \delta_i(K) \quad \text{if} \\
&\max\left(\frac{\frac{1}{2}\left[\delta_{i-1}^2 + \delta_i^2\right] - \omega}{\alpha\sigma_t^{(K)2}} - \frac{\beta}{\alpha}, 0\right) \leq \\
&\left[q_1(z_t - \theta)^+ + q_2(z_t - \theta)^- + (1 - q_1 - q_2)|\xi_t|\right]^2 < \\
&\max\left(\frac{\frac{1}{2}\left[\delta_i^2 + \delta_{i+1}^2\right] - \omega}{\alpha\sigma_t^{(K)2}} - \frac{\beta}{\alpha}, 0\right).
\end{aligned} \tag{A.1}$$

with $\begin{bmatrix} z_{t+1}^* \\ \xi_{t+1} \end{bmatrix} |\Psi_t \overset{\mathbb{P}}{\sim} \mathbb{N}\left(\mathbf{0}_{2\times 1}, \boldsymbol{I}_{2\times 2}\right)$.

Assume that the partition condition (I.3.16) is satisfied and, without loss of generality, that $\delta_1 < \delta_2 < \ldots < \delta_K$. Now, by construction, the 'max' functions in (A.1) can be dropped without loss of generality, since if the lower bound was non-positive for all $i > 1$, the partition would call for $\sigma_{t+1}^{(K)} = \delta_1$ and analogously, if the upper bound was non-positive for all $i < K$, the partition would imply $\sigma_{t+1}^{(K)} = \delta_K$.

Simplifying and rearranging then leads to

$$\begin{aligned} \frac{1}{2}[\delta_{i-1}^2 - \delta_i^2] \leq \omega + \beta\sigma_t^{(K)2} + \alpha\sigma_t^{(K)2}\big[q_1(z_t-\theta)^+ + q_2(z_t-\theta)^- \\ + (1-q_1-q_2)|\xi_t|\big]^2 - \delta_i^2 < \frac{1}{2}[\delta_{i+1}^2 - \delta_i^2]. \end{aligned} \tag{A.2}$$

Assuming that the partition condition holds, the gap between two adjacent volatility levels δ_i and δ_{i+1} approaches zero as the number of regimes approach infinity. Hence, the volatility process can equivalently be written as

$$\sigma_{t+1}^{(K)2} = \omega + \beta\sigma_t^{(K)2} + \alpha\sigma_t^{(K)2}\left[q_1(z_t-\theta)^+ + q_2(z_t-\theta)^- + (1-q_1-q_2)|\xi_t|\right]^2 + h(K) \tag{A.3}$$

where $h(K) \to 0$ as $K \to \infty$. Hence, for a fixed $\sigma_t^{(K)}$ and $\omega \geq 0, \alpha \geq 0, \beta \geq 0$, (A.2) and (A.3) imply that $\sigma_{t+1}^{(K)}$ converges *almost surely* in $\mathbb{P}$ to

$$\sigma_{t+1}^{(K)2} = \omega + \beta\sigma_t^{(K)2} + \alpha\sigma_t^{(K)2}\left[q_1(z_t-\theta)^+ + q_2(z_t-\theta)^- + (1-q_1-q_2)|\xi_t|\right]^2,$$

which implies that $\sigma_{t+1}^{(K)2} \overset{a.s.}{\to} \sigma_{t+1}^2$ leading to $S_t^{(K)} \overset{a.s.}{\to} S_t$ in $\mathbb{P}$ as $K \to \infty$. Now, assuming that $S_0^{(K)} = S_0$ and $\sigma_1^{(K)2} = \sigma_1^2$ are known, the same procedure can be used over the full period $[0, T]$, implying that (I.3.17) converges *almost surely* in $\mathbb{P}$ to (I.3.27). $\square$

B Figures and Tables

FIGURE B.1: S&P 500 closing levels from January 01, 2011 till July 31, 2014.

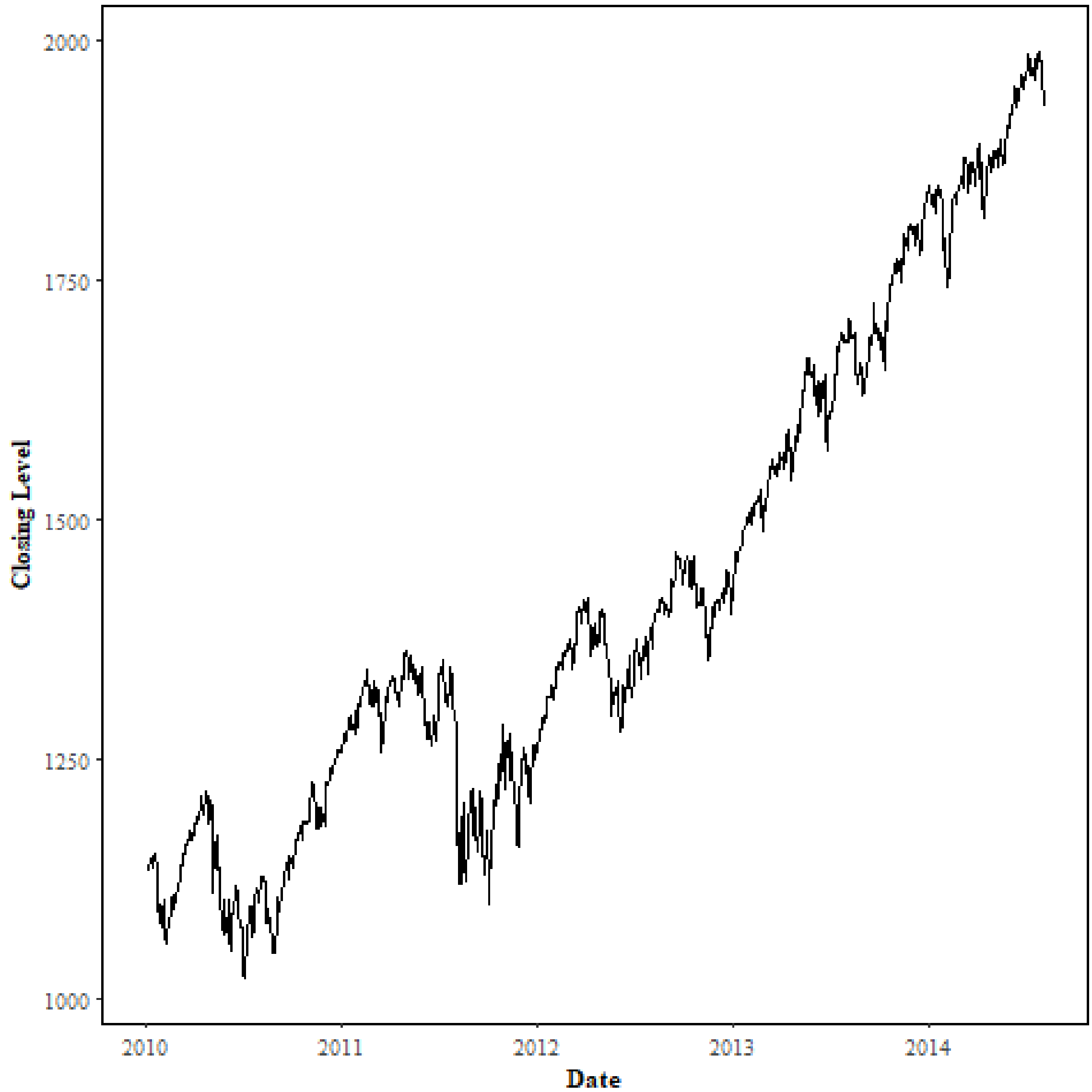

FIGURE B.2: S&P 500 daily log returns from January 01, 2011 till July 31, 2014.

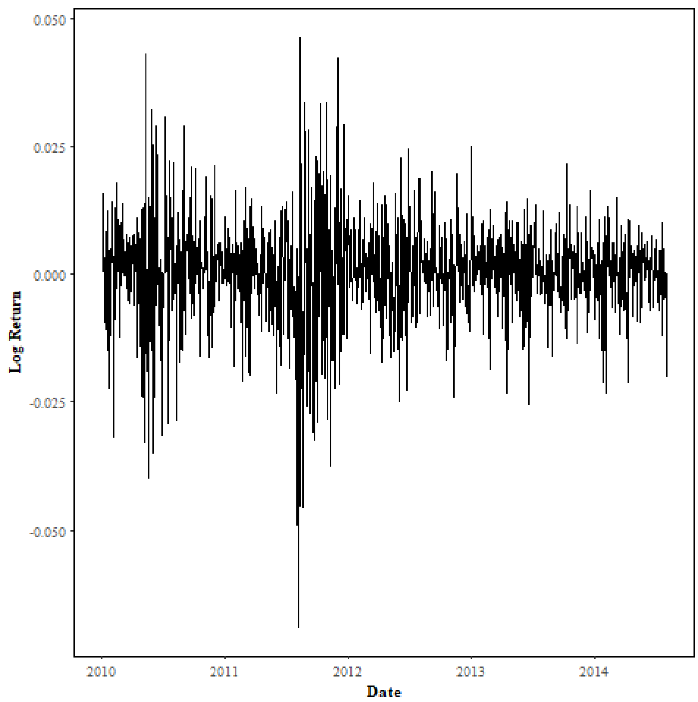

FIGURE B.3: S&P 500 daily absolute log returns from January 01, 2011 till July 31, 2014.

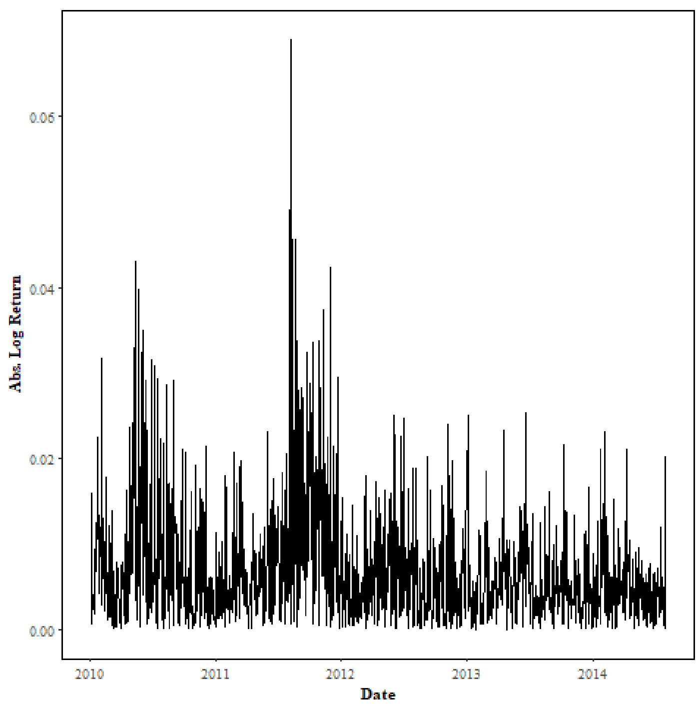

FIGURE B.4: NAGARCH-P portfolio returns for different risk aversion levels without transaction costs.

FIGURE B.5: NAGARCH-P portfolio returns for different risk aversion levels with transaction costs.

FIGURE B.6: NAGARCH-Q portfolio returns for different risk aversion levels without transaction costs.

FIGURE B.7: NAGARCH-Q portfolio returns for different risk aversion levels with transaction costs.

FIGURE B.8: U-MSM-2 portfolio returns for different risk aversion levels without transaction costs.

FIGURE B.9: U-MSM-2 portfolio returns for different risk aversion levels with transaction costs.

FIGURE B.10: B-MSM-11 portfolio returns for different risk aversion levels without transaction costs.

FIGURE B.11: B-MSM-11 portfolio returns for different risk aversion levels with transaction costs.

FIGURE B.12: B-MSM-Inf portfolio returns for different risk aversion levels without transaction costs.

FIGURE B.13: B-MSM-Inf portfolio returns for different risk aversion levels with transaction costs.

FIGURE B.14: ABS portfolio returns for different risk aversion levels without transaction costs.

FIGURE B.15: ABS portfolio returns for different risk aversion levels with transaction costs.

FIGURE B.16: VIX portfolio returns for different risk aversion levels without transaction costs.

FIGURE B.17: VIX portfolio returns for different risk aversion levels with transaction costs.

FIGURE B.18: S&P 500 (blue, left) and VIX (red, right) between January 01, 2010 and July 31, 2014.

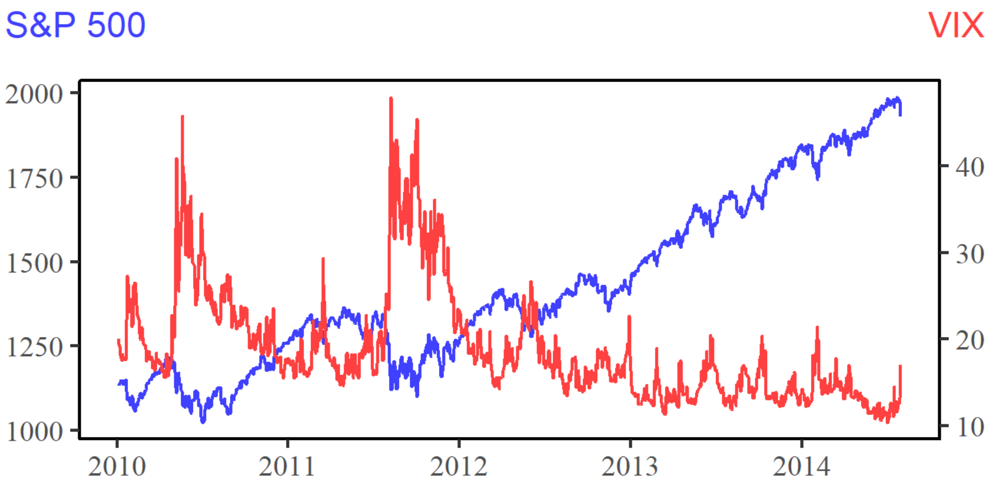

FIGURE B.19: Implied Volatility Plot: ABS vs BS (01/01/2010, 10 dtm).

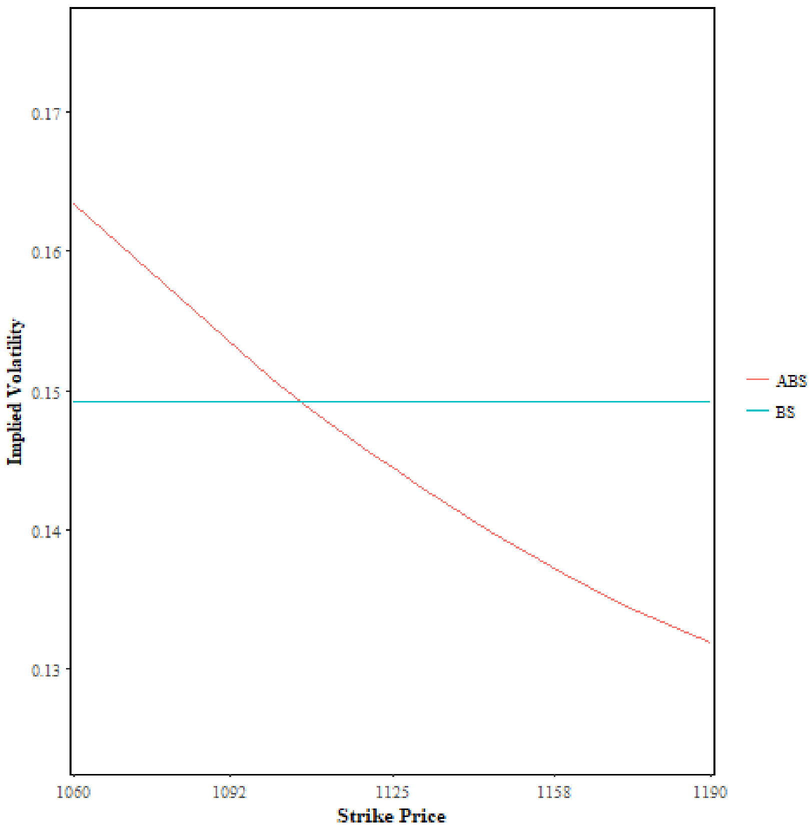

TABLE B.1: Number of contracts across moneyness and maturity.

Maturity	Moneyness	In-sample	Week 1	Week 2	Week 3	Total
6-50 days Short term contracts	$M_t < -0.075$	23	18	21	23	85
	$-0.075 \leq M_t < -0.050$	56	50	56	54	216
	$-0.050 \leq M_t < -0.025$	176	143	187	144	650
	$-0.025 \leq M_t < 0.000$	561	559	557	525	2202
	$0.000 \leq M_t < 0.025$	1084	1111	1139	1041	4375
	$0.025 \leq M_t < 0.050$	848	827	810	779	3264
	$0.050 \leq M_t < 0.075$	371	375	395	366	1507
	$M_t \geq 0.075$	154	164	165	144	627
51-100 days Long term contracts	$M_t < -0.075$	13	16	6	11	46
	$-0.075 \leq M_t < -0.050$	19	25	21	25	90
	$-0.050 \leq M_t < -0.025$	48	49	51	51	199
	$-0.025 \leq M_t < 0.000$	197	204	201	225	827
	$0.000 \leq M_t < 0.025$	341	307	349	315	1312
	$0.025 \leq M_t < 0.050$	319	294	285	285	1183
	$0.050 \leq M_t < 0.075$	257	235	283	253	1028
	$M_t \geq 0.075$	170	160	167	172	669
Total		4637	4537	4693	4413	18280

The table shows the number of call contracts across different maturity and moneyness brackets available in the data set. Moneyness is defined as $M_t := K/F_t - 1$. The data is shown for the in-sample weeks and the three following out-of-sample weeks between January 01, 2010 and July 31, 2014.

TABLE B.2: Average quoted mid price across moneyness and maturity.

Maturity	Moneyness	In-sample	Week 1	Week 2	Week 3	All
6-50 days Short term contracts	$M_t < -0.075$	120.19	120.60	136.37	121.05	124.51
	$-0.075 \leq M_t < -0.050$	90.47	94.70	90.86	92.61	92.08
	$-0.050 \leq M_t < -0.025$	60.59	59.98	59.99	61.56	60.50
	$-0.025 \leq M_t < 0.000$	28.90	28.44	28.93	29.77	29.00
	$0.000 \leq M_t < 0.025$	11.52	11.20	10.98	11.50	11.29
	$0.025 \leq M_t < 0.050$	3.67	3.78	3.69	4.01	3.78
	$0.050 \leq M_t < 0.075$	2.13	2.33	2.00	2.11	2.14
	$M_t \geq 0.075$	1.61	1.38	1.21	1.57	1.44
51-100 days Long term contracts	$M_t < -0.075$	136.03	138.22	147.10	123.85	123.85
	$-0.075 \leq M_t < -0.050$	100.49	104.25	103.00	105.82	105.82
	$-0.050 \leq M_t < -0.025$	74.29	74.85	71.32	74.55	74.55
	$-0.025 \leq M_t < 0.000$	47.04	47.30	47.22	46.39	46.39
	$0.000 \leq M_t < 0.025$	30.23	30.85	31.21	30.49	30.49
	$0.025 \leq M_t < 0.050$	13.60	14.83	15.06	14.24	14.24
	$0.050 \leq M_t < 0.075$	6.43	7.28	7.34	6.05	6.05
	$M_t \geq 0.075$	3.99	4.94	4.15	3.58	3.58

The table shows the average recorded market prices for all call contracts across different maturity and moneyness brackets in the data set. Moneyness is defined as $M_t := K/F_t - 1$. The data is shown for the in-sample weeks and the three following out-of-sample weeks between January 01, 2010 and July 31, 2014.

TABLE B.3: Parameter estimates

Models	Parameters					
NAGARCH-P	ω	α	β	λ	θ		
Mean	$2.05E-06$	0.0724	0.8473	0.0241	0.9844		
StDev	$5.64E-08$	0.0020	0.0009	0.0033	0.0302		
NAGARCH-Q	ω	α	β	θ^*			
Mean	$2.04E-06$	0.0863	0.7875	1.0676			
StDev	$1.84E-06$	0.0692	0.1690	0.5595			
U-MSM-2	δ_1	δ_2	p_{11}	p_{22}			
Mean	0.0076	0.0080	0.9539	0.9940			
StDev	0.0030	0.0015	0.1090	0.0301			
B-MSM-11	ω	α	β	θ^*	q_1	q_2	$\bar{\sigma}_{t,T}$
Mean	$1.07E-07$	0.2034	0.8215	0.6280	0.0606	0.0216	0.0076
StDev	$1.96E-06$	0.2736	0.2835	3.5376	0.1193	0.0565	0.0023
B-MSM-Inf	ω	α	β	θ^*	q_1	q_2	σ_t
Mean	$8.06E-06$	0.1465	0.2897	3.7709	0.4442	0.5479	0.0069
StDev	$4.88E-06$	0.0314	0.2445	0.8919	0.1064	0.1058	0.0059
ABS	β_0	β_1	β_2	β_3	β_4	β_5	
Mean	0.0406	$-4.7239E-05$	$1.6427E-08$	$1.9893E-04$	$-4.1895E-08$	$-1.2453E-07$	
StDev	$4.88E-06$	0.0314	0.2445	0.8919	0.1064	0.1058	0.0059

The table presents the parameter estimates obtained during the monthly estimation procedure. Shown are the sample means and standard deviations across all 60 weekly estimations.

TABLE B.4: Mean square errors in-sample and out-of-sample

Models	In-sample	Week 1	Week 2	Week 3
NAGARCH-Q	2.3784 (0.9010)	6.6483 (3.8530)	11.8423 (9.2123)	13.0309 (10.2575)
U-MSM-2	6.4502 (1.8569)	7.1739 (4.1074)	11.4403 (8.5114)	15.5383 (14.5939)
B-MSM-11	1.9733 (0.8699)	10.2937 (3.3326)	15.4064 (13.0216)	22.5833 (14.0324)
B-MSM-Inf	2.2646 (0.8330)	7.0041 (3.2715)	12.2179 (10.8232)	13.9965 (11.8431)

The sample mean and standard deviations of the model's MSE values are presented, both in-sample and out-of-sample. U-MSM-2 and B-MSM-11 use the average state mean square errors.

TABLE B.5: Optimal portfolio allocations

Models	$\kappa = 2$	$\kappa = 4$	$\kappa = 8$
NAGARCH-Q	1.0509 (0.3606)	0.5279 (0.1826)	0.2646 (0.0920)
U-MSM-2	1.0534 (0.3690)	0.5291 (0.1870)	0.2653 (0.0942)
B-MSM-11	1.0044 (0.5922)	0.5389 (0.3818)	0.2697 (0.1901)
B-MSM-Inf	1.3969 (0.5899)	0.9597 (0.6491)	0.7008 (0.6988)

Displayed are the sample means and standard errors of the weight allocations to the risky asset for each strategy and κ.

TABLE B.6: Portfolio returns

Models	Trading costs	$\kappa = 2$	$\kappa = 4$	$\kappa = 8$
NAGARCH-P	Yes	17.69%	9.48%	4.82%
	No	19.46%	9.74%	4.95%
NAGARCH-Q	Yes	18.15%	9.73%	4.95%
	No	19.97%	10.00%	5.08%
U-MSM-2	Yes	14.18%	7.06%	3.77%
	No	16.86%	7.81%	4.02%
B-MSM-11	Yes	17.93%	10.48%	5.30%
	No	21.60%	10.69%	5.38%
B-MSM-Inf	Yes	19.26%	13.74%	9.85%
	No	23.74%	16.48%	12.33%

The table presents the annualized portfolio returns for several risk aversion parameters with and without trading costs. As a benchmark, the S&P 500 index achieved an annualized return of 16.82% over the same period.

TABLE B.7: Portfolio Sharpe ratios

Models	Trading costs	$\kappa = 2$	$\kappa = 4$	$\kappa = 8$
NAGARCH-P	Yes	0.8196	0.7317	0.4241
	No	0.9189	0.7604	0.4528
NAGARCH-Q	Yes	0.8387	0.7546	0.4490
	No	0.9381	0.7833	0.4776
U-MSM-2	Yes	0.5808	0.3878	0.1605
	No	0.7203	0.4585	0.2063
B-MSM-11	Yes	0.5942	0.6056	0.3821
	No	0.7183	0.6228	0.3974
B-MSM-Inf	Yes	0.6687	0.6171	0.4719
	No	0.8502	0.7738	0.6400

The table presents the Sharpe ratios for each strategy for several risk aversion parameters and with and without trading costs. The risk-free rate is chosen as the mean of the short-ends of the LIBOR curve during the period, which amounts to 2.92% annually and $3.24E - 06$ daily. As a benchmark, a buy-and-hold strategy using the S&P 500 would have achieved a Sharpe ratio of 0.7069 given this interest rate.

TABLE B.8: Certainty equivalent returns

Models	Trading costs	$\kappa = 2$	$\kappa = 4$	$\kappa = 8$
NAGARCH-P	Yes	0.23	0.13	-0.11
	No	0.28	0.24	-0.08
NAGARCH-Q	Yes	0.24	0.17	-0.09
	No	0.32	0.25	-0.06
U-MSM-2	Yes	0.11	-0.22	-0.37
	No	0.13	0.01	-0.31
B-MSM-11	Yes	0.21	-0.12	-0.15
	No	0.22	0.20	-0.13
B-MSM-Inf	Yes	0.30	0.16	-0.10
	No	0.42	0.22	-0.06

The table presents the excess annualized certainty equivalent returns (CER) for each strategy for several risk aversion parameters, with and without trading costs, over the S&P 500 index during the studied period. Results are presented in basis points (bp). As a benchmark, a buy-and-hold strategy using the S&P 500 would have achieved CER's of 113bp, 79bp and 12bp for the different risk aversion levels, respectively.

TABLE B.9: Mean square errors in-sample and out-of-sample for ABS vs BS

	In-sample	Week 1	Week 2	Week 3
ABS	1.9782	8.7026	13.0883	16.3166
	(2.8389)	(14.2423)	(16.8888)	(20.5249)
BS	9.9463	16.7210	20.8447	31.9308
	(4.6135)	(13.8585)	(18.5099)	(65.7863)

The sample mean and standard deviations of the ABS and BS models' MSE values are presented, both in-sample and for each of the three out-of-sample weeks.

TABLE B.10: ABS Portfolio Strategy

Measure	Trading costs	$\kappa = 2$	$\kappa = 4$	$\kappa = 8$	S&P 500
Portfolio Weights		1.4384	0.7357	0.3678	1.0000
		(0.4139)	(0.2383)	(0.1191)	(0.0000)
Annualized Returns	Yes	18.67%	11.30%	5.78%	16.82%
	No	22.88%	11.77%	5.98%	
SR	Yes	0.6369	0.6654	0.4535	0.7069
	No	0.8072	0.7032	0.4862	
CER	Yes	1.09%	0.71%	0.36%	$*$
	No	1.46%	0.75%	0.38%	

The table presents the average inferred portfolio weights of ABS, with its standard errors in parentheses, as well as its annualized portfolio returns, SR and CER for several risk aversion parameters with and without trading costs.
$*$ As a benchmark, the performance of a S&P 500 buy-and-hold strategy is shown, which accomplishes CER's of 1.13%, 0.79% and 0.12% for the different risk aversion levels, respectively.

TABLE B.11: VIX Portfolio Strategy

Measure	Trading costs	$\kappa = 2$	$\kappa = 4$	$\kappa = 8$	S&P 500
Portfolio Weights		1.0816	0.5408	0.2704	1.0000
		(0.3370)	(0.1685)	(0.0843)	(0.0000)
Annualized Returns	Yes	15.86%	8.65%	4.41%	16.82%
	No	17.56%	8.88%	4.54%	
SR	Yes	0.7127	0.6314	0.3269	0.7069
	No	0.8066	0.6562	0.3552	
CER	Yes	1.09%	0.61%	0.31%	$*$
	No	1.24%	0.63%	0.32%	

The table presents the average inferred portfolio weights of the VIX strategy, with its standard errors in parentheses, as well as its annualized portfolio returns, SR and CER for several risk aversion parameters with and without trading costs.
$*$ As a benchmark, the performance of a S&P 500 buy-and-hold strategy is shown, which accomplishes CER's of 1.13%, 0.79% and 0.12% for the different risk aversion levels, respectively.

TABLE B.12: ABS vs BS parameter estimates

	Parameters					
ABS	β_0	β_1	β_2	β_3	β_4	β_5
	$4.0641E-02$	$-4.7239E-05$	$1.6427E-08$	$1.9893E-04$	$-4.1895E-08$	$-1.2453E-07$
	$(2.7063E-01)$	$(3.2080E-04)$	$(9.8771E-08)$	$(6.7783E-04)$	$(5.4320E-07)$	$(4.6187E-07)$
BS	β_0					
	$7.6230E-03$					
	$(2.5537E-03)$					

The table presents the ABS and BS parameter estimates obtained during the monthly estimation procedure. Shown are the sample means and standard deviations across all 60 weekly estimations.