MSc Stochastics and Financial Mathematics

Master thesis

Towards a long maturity pricing kernel under generalised Heston & Nandi (2000) GARCH

by

Minne Marijnen

6th July 2020

Supervisor: prof. dr. Peter Spreij (UvA)

Ortec Finance Supervisor: Pieter Kloek (OF)

Second examiner: dr. Dennis Dobler (VU)

Department of Mathematics
Faculty of Sciences

VU UNIVERSITY AMSTERDAM

ORTEC FINANCE

Abstract

In order to value options, we rely on models that are calibrated under the risk-neutral measure $\mathbb{Q}$. However, in a long maturity context, products exist that include optionality, and therefore need valuation under $\mathbb{Q}$, but for which no market prices are available. The generalised version by Christoffersen et al. (2013) of the Heston & Nandi (2000) GARCH model encompasses both a transformation from physical measure $\mathbb{P}$ to $\mathbb{Q}$, and a closed-form option pricing formula. It is demonstrated that a time series approach by this model, particularly when enhanced with calibration against the economical scenarios by Ortec Finance, can be fruitful for valuing near the money options with increasing maturity, while it is unable to match the severe strike price skewness in market implied volatilities. A near linear increase in the discrepancy between $\mathbb{P}$ and $\mathbb{Q}$ with respect to maturity is observed.

Acknowledgements

This thesis was written for the fulfilment of the Master of Science program 'Stochastics and Financial Mathematics' at the Vrije Universiteit (VU) and the Universiteit van Amsterdam (UvA). The research was conducted at Ortec Finance, under supervision of Pieter Kloek. I am thankful to Ortec Finance for this opportunity, and I would hereby like to thank Erik Hennink and Pieter for their input. In particular I would like to thank Pieter for his continuous dedication and enthusiasm with regard to this project.

My academic supervisor was Peter Spreij (UvA). I also want to thank Peter for his feedback and suggestions throughout the project.

Contents

List of Figures

List of Tables

1. Introduction

1.1. Background

In the derivation of the famous Black-Scholes formula it is assumed that stock returns are log-normally distributed with constant volatility[a]. An array of empirical studies, such as Rubinstein (1985), show that these assumptions do not hold. As a result the classical Black-Scholes model ignores stylised facts of asset returns, notably:

- Asset returns are leptokurtic, or fat-tailed. As was already documented by the early work in Mandelbrot (1963), Fama (1963), Fama (1965), it follows that the normality assumption of log returns does not hold.
- Volatility clustering: High/low deviations in stock returns tend to cluster. This characteristic violates the assumption that volatility is constant, and is known as conditional heteroskedasticity.
- Leverage effect: Asset returns are negatively correlated with volatility. This is recognised and described as leverage effect by Black (1976).

As a consequence of the first two properties, the implied volatility[b] is higher for deeper in the money/out of the money options, as is empirically validated in the option market. This phenomenon is known as the volatility smile. The third property leads to skewness in this implied volatility smile, which creates a so-called volatility smirk. Cont (2001) provides an extensive overview of stylised facts in asset returns.

1.2. Stochastic volatility models

A number of stochastic volatility models have been developed as extensions of the Black-Scholes model in order to accommodate these stylised facts. Stochastic volatility models can be branched into (continuous time) stochastic volatility models and (discrete) time series models. Popular examples of the first category are Hull & White (1987) and Heston (1993). Continuous SV-models tend to have desirable characteristics for option pricing, since 'They manage to reproduce empirical regularities displayed by the risk-neutral density implied in option quotations' (Moyaert & Petitjean 2011). A disadvantage is that they can be computationally expensive, and often a closed-form option

[a] In finance, the term volatility is used for the conditional variance of asset returns.
[b] See section 4.5

pricing formula is no longer available. Heston (1993) derives an affine[c] semi-closed-form option pricing formula , thereby making it a popular choice in its category.

Since conceived by Engle (1982) (as ARCH) and generalised by Bollerslev (1986) GARCH models h ave gained great success in the modelling of heteroskedasticity. An important advantage of GARCH models with respect to continuous time models is that volatility is readily observable from historical realisations, allowing option pricing based on only the history of the underlying asset, whereas in the continuous time setting it is impossible to recover volatility directly from discrete historical observations.

A popular version of the GARCH-model was introduced by Heston & Nandi (2000) and generalised by Christoffersen et al. (2013). This model will hereafter be referred to as the HNG-model. An important advantage of the HNG-model is that it entails a closed-form option pricing formula. This is in contrast to many other GARCH type models that rely on Monte Carlo simulation methods to compute option prices. The closed-form option pricing formula is derived by stipulating affineness. Heston & Nandi (2000) show that the HNG-model contains the Heston (1993) model as a continuous limit. A more detailed version of this derivation is given in appendix D.3. As shown by Su et al. (2010), the HNG-model is able to model heteroskedasticity in returns well, and also thanks to its closed-form formula for option pricing it is 'very popular in option pricing literature' (Wang et al. 2017).

1.3. Embedded options

This thesis concerns the issue of valuation of embedded options, which is relevant for life insurance companies. For these institutions, not only the risk and return on asset classes needs to be simulated, but also liabilities. A typical balance sheet of a life insurance company will contain contracts with embedded guarantees. Examples are unit-linked (UL), profit sharing and variable annuity products. These contracts may bring a profit to the holder (buyer), but never a loss, depending on underlying variables such as interest rates and equity. Therefore, in a mathematical sense, these are option contracts, and these embedded guarantees are therefore denoted as embedded options.

It is pointed out by Flint et al. (2015) that "There exists a noticeable dearth in long-term volatility forecasting and analysis, while this should be a concern given the large quantity of life policies". In the literature on models such as the HNG-model options with maturity of more than 1 year are seldom considered. In practice (embedded) options on a much longer (>10 year) range require

[c] A (pay-off) process X is called affine if its conditional characteristic function is exponential affine in X_t for all $t \leq T$, that is, there exist $\mathbb{C}$ and $\mathbb{C}^d$-valued functions $A(t,u)$, $B(t,u)$ with jointly continuous t-derivatives such that

$$\mathbb{E}[e^{u^T X(T)} | \mathcal{F}_t] = \exp\{A(T-t,u) + B(T-t,u)^T X(t)\}$$

for all $u \in i\mathbb{R}^d$, $t \leq T$.

apt modelling. Flint et al. (2015) makes an effort to fill this dearth, but does not implement the HNG-model. Flint et al. (2015) justify this by stating that GARCH is not intended for long-term. However, Moyaert & Petitjean (2011) find that "Medium or long maturity options are better treated by the HNG-model, given their flatter volatilities", and amongst others Lassance & Vrins (2018) confirm that HNG is relatively well suited to longer maturities.

1.4. The pricing kernel

What remains a significant issue in the context of option valuation is the consistency between the distribution as implied by market option prices (known as the 'risk-neutral distribution', hereafter denoted as 'distribution under the risk-neutral measure $\mathbb{Q}$') and distribution of the underlying asset ('distribution under $\mathbb{P}$'). It is argued by Bates (1996) that this is in fact 'the central empirical issue' in option research. As a consequence of the first fundamental theorem of asset pricing, absence of market arbitrage induces that a so-called pricing kernel, that is, a Radon-Nikodym derivative $\frac{\mathrm{d}\mathbb{P}}{\mathrm{d}\mathbb{Q}}$ that relates $\mathbb{P}$-distribution to $\mathbb{Q}$-distribution, exists. This relation is also commonly known as the State-Price Density (SPD). The issue of finding a suitable pricing kernel can only be scrutinised under assumption of a particular distributional hypothesis. See Bates (1996) for an overview of empirical literature in this area. As is common practice in related empirical studies, Heston & Nandi (2000) use a double calibration approach to circumvent the consistency issue between $\mathbb{P}$- and $\mathbb{Q}$-distribution. Heston & Nandi (2000) fit parameters of the HNG-model under $\mathbb{P}$ by maximum likelihood maximisation with respect to historical data of the underlying time series, and thereafter a second calibration round of these model parameters is performed under $\mathbb{Q}$, by a non-linear least squares procedure with respect to market option prices.

Christoffersen et al. (2013) generalise the HNG-model by introducing a variance-dependent pricing kernel (a relation between $\mathbb{P}$ and $\mathbb{Q}$ distribution), that implies a simple mapping from $\mathbb{P}$- to $\mathbb{Q}$-parameters. This mapping depends on a single parameter ξ. However, Christoffersen et al. (2013) perform a joint estimation procedure that allows the full model to be simultaneously estimated under $\mathbb{P}$ and $\mathbb{Q}$, as opposed to estimating parameters under $\mathbb{P}$ and mapping these to $\mathbb{Q}$-parameters. They add a brief note on a so-called sequential parameter estimation method, which truthfully bridges the $\mathbb{Q}$- and $\mathbb{P}$-worlds. This approach becomes of particular interest in a setting where no option data is available for calibration under $\mathbb{Q}$, and will be further explained below.

1.5. Research topic

In this thesis, it is examined whether the generalised HNG-model is well suited to be used as a reference for the valuation of long maturity embedded options. By absence of market data, the model cannot be calibrated under $\mathbb{Q}$, as no

market data for equity options with maturity of more than ten years is available. For this reason, the aim is to specify an adequate pricing kernel within the HNG framework that can reasonably be assumed to hold for maturity beyond 10 years. A candidate is the variance-dependent pricing kernel by Christoffersen et al. (2013). To scrutinise if this is a suitable candidate, data of market option prices up to ten years is used. The following approach is adopted: First the HNG-model is calibrated under $\mathbb{P}$. Option prices as computed with these $\mathbb{P}$-parameters are compared with market option prices, and the severe discrepancy between distribution under $\mathbb{P}$ and $\mathbb{Q}$ is confirmed. Subsequently an extra parameter ξ, that describes the mapping from $\mathbb{Q}$ to $\mathbb{P}$, is estimated under $\mathbb{Q}$, against market option prices with maturity varying from one up to ten years. This method is denoted as the sequential parameter estimation method. It is found that by this method the level of the market implied volatility is reached, but the implied volatility skew under this model is less skewed than it is from market options prices.

Following Christoffersen et al. (2013), a benchmark is set by calibrating the HNG-parameters under $\mathbb{Q}$, thereby completely ignoring the $\mathbb{Q}/\mathbb{P}$-mapping that is described by the variance dependent pricing kernel. This method will be denoted as the ad-hoc calibration method. In the literature it was found that this $\mathbb{Q}$-calibrated HNG-model is able to perform well in modelling option prices. This finding is indeed confirmed to remain valid on a long maturity horizon (up to ten years) as well. The error margins, both in- and out of sample, of the sequential and ad-hoc models are compared, in order to evaluate if the estimated $\mathbb{Q}/\mathbb{P}$ mapping was adequate. The evolution of the shape of this mapping for maturities increasing from 1 year to 10 years is evaluated. It is found that a nearly linear pattern is seen between the estimates of this mapping parameter and maturity T, indicating a positive correlation between risk averseness and option maturity.

The implied volatility skew based on the daily formulation of the HNG-model is seen to lose skewness for a longer maturity horizon. Consequently, experiments are performed with reformulation of the model with longer time steps. The yearly formulation of the sequential model is shown to better reflect the market skewness than the daily version.

Subsequently, an adaptation of the sequential estimation method is performed, in which a set of economic scenarios is used as pseudo-data. The $\mathbb{P}$-calibration step is performed again by fitting the model against scenarios from the Ortec Finance Scenario set (OFS). These are provided by Ortec Finance, and are generated by an advanced econometric model. See section 2.1 for more background on these scenarios. It is evaluated whether this method of estimation against forward-looking data enhances the ability of the model to model and predict option prices. Without any $\mathbb{P}/\mathbb{Q}$-mapping, so under $\mathbb{P}$, this version of the model has more skewness than the regular historically calibrated model. After sequential estimation the difference is smaller, but remains present.

1.6. Layout

The data that is used includes market data and scenarios from the (real-world) Ortec Finance Scenario set (OFS). Both are described in section 2. An introduction of the general class of GARCH models is given in section 3. The derivation and implementation results of the basic HNG-model and generalisations are outlined in the section 4. Results are displayed in section 5.

2. Data

S&P 500 historical data is used. In the Bloomberg terminal the following description is given: "The S&P 500 is widely regarded as the best single gauge of large-cap U.S. equities and serves as the foundation for a wide range of investment products. The index includes 500 leading companies and captures approximately 80% coverage of available market capitalisation." Consequently, this data series is the standard in option pricing literature, see for example Shu & Zhang (2003), who state: "The S&P 500 index options have the largest open interest and the largest trading volume in all options traded on the Chicago Board Options Exchange (CBOE). Unlike S&P 100 index options, S&P 500 index options are European style and do not have the problem of early exercise".

This data series is given in daily closing prices. It is available from 1927-12-30. Data up until 2020-02-03 is used. The data set contains index values of 23132 trading days. Different subsets of this data range are used.

2.1. Ortec Finance Scenario set

Ortec Finance Scenario set (OFS) is generated by the advanced econometric model that is used by Ortec Finance. This model, that is build on a factor based approach, is called the Dynamic Scenario Generator. See Steehouwer (2019) for a detailed outline of the OFS approach.

A set of 2000 scenarios containing monthly data points on a ten year horizon for US equity total returns and 3-month swap rate is used. These scenarios are generated from the December 2018 calibration of the Ortec Finance model. See figure 2.1 for a fan-chart of the US equity scenarios.

2.2. ICE Data Services implied volatility data

ICE Data Services (Superderivatives) provides Ortec Finance with market option data. In this thesis the S&P 500 implied volatility surfaces are used. These are provided monthly and contain implied volatilities for maturities between 1 week and 10 years, for relative moneyness ranging from -50% to $+50\%$. This data is available between 2017-07 and 2020-02, with a total of 3388 implied volatilities. The term structure of dividend yield for each of the maturities is provided in the same monthly dataset.

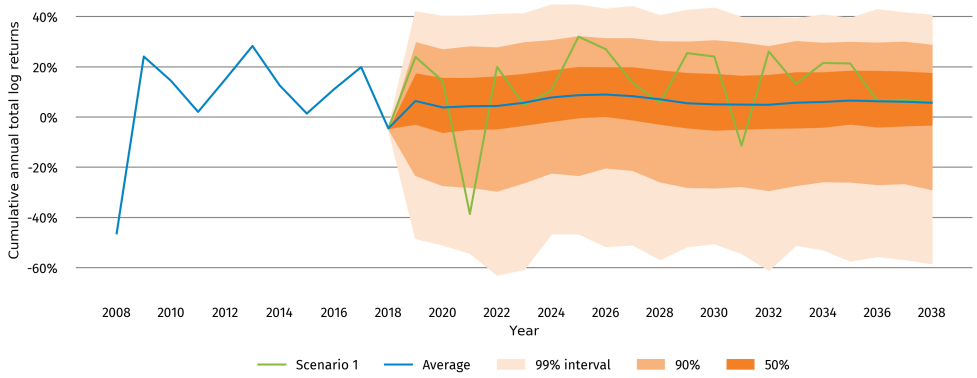

Figure 2.1. A fan-chart of the US equity scenarios from the OFS
The Ortec Finance scenarios for US equity total return on a 20 years horizon is displayed in a fan-chart with confidence bands. These scenarios are generated by the December 2018 scenario set, and are given in annual cumulative log returns format. Ten years of history is displayed as well.

2.3. Historical dividends

As the HNG-model is formulated for total returns, the historical S&P 500 pricing return data is converted to total return format. For this purpose S&P 500 monthly dividend yield data is used.

2.4. Interest rates

As historical interest rates, the 3-month treasury bill yield is used. This is gathered from Bloomberg. For the valuation of options the 3-month swap rate is used. This is available for maturities between 1 week and 60 years in the interest rate yield-curve data, that is monthly provided by ICE.

3. GARCH models

The hypothesis that log returns of assets follow a white noise process[a] was still widely accepted until the 1970s, underlying the Black-Scholes option pricing formula (see (4.15)). In the 1980s, Engle (1982) generalised this assumption with the introduction of a model that allowed for volatility clustering, namely the Auto Regressive Conditional Heteroskedasticity (ARCH) model. This model was further generalised with the introduction of the GARCH model in Bollerslev (1986):

Definition 1. Let $(\mathcal{F}_t)_{t\in\mathbb{Z}}$ be a filtration. A GARCH(p,q) process is an adapted process $(X_t)_{t\in\mathbb{Z}}$ where X_t, is normally distributed conditional on $\mathcal{F}_{t-1}$ with zero mean and variance h_t. The conditional variance $h_t = \mathbb{E}[X_t^2|\mathcal{F}_{t-1}]$ follows the process

$$h_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i X_{t-i}^2 + \sum_{i=1}^{p} \beta_i h_{t-i} \tag{3.1}$$

where $p \geq 0$, $q > 0$, $\alpha_0 > 0$, $\alpha_i \geq 0$, $i \in [1, .., q]$, $\beta_i \geq 0$, $i \in [1, ..., p]$.

For $p = 0$ the ARCH process is recovered. Francq & Zakoïan (2010) show that the following is a necessary and sufficient condition for second-order stationarity of the GARCH-process:

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

where the quantity $\sum_{i=1}^{q} \alpha_i + \sum_{j=1}^{p} \beta_j$ is known as the persistence, which characterises the level of decay after a shock. A persistence that is close to 1 indicates that the process recovers slowly from a shock, while a near-zero persistence indicates rapid mean-reversion. The process explodes if it is non-stationary, i.e. if the persistence is larger than 1.

The assumption that for all t, X_t is conditionally normally distributed was adopted in the original papers by Engle (1982) and Bollerslev (1986), but may well be replaced by assuming that the X_t follow another conditional distribution with zero expectation and variance h_t. See for example Bai et al. (2003) for a kurtosis-related discussion with varying innovation distribution of the GARCH process.

[a] A white noise process is a stationary time series X_t with mean zero and with $\mathbb{E}[X_t^2] = \sigma^2$, $\mathbb{E}[X_t X_s] = 0$ for $t \neq s$.

In contrast to a random walk[b] model, the standard GARCH-model is able to reproduce two of the prominent stylised facts[c] in asset returns that were mentioned in the introduction of this thesis, namely volatility clustering and leptokurtosis, or fatness of tails. Volatility clustering is the property that periods of higher volatility tend to be alternated with periods of lower volatility. This is equivalent to the property that auto-covariances are positive. See Kumar & Dhankar (2010) for empirical proof of volatility clustering in US equity returns. Francq & Zakoïan (2010) show that positivity[d] of the auto-covariances holds for the GARCH-process as well, so that this model is able to reproduce this stylised fact of asset returns.

Another stylised fact of asset returns is that the tails of the distribution are fatter than the tails of a normal distribution, i.e. asset returns are leptokurtic (see e.g. Bollerslev (1987) for empirical validation of this stylised fact). Francq & Zakoïan (2010) show that the excess kurtosis of a GARCH process[e] (3.1) is strictly positive given that the process admits moments up to order 4.

This shows that of the three stylised facts of asset returns that were mentioned in the introduction, two are captured by the standard GARCH model. The third stylised fact that was mentioned is the leverage effect, i.e. negative correlation of volatility and innovation. In order to find a model that is able to capture this effect as well, a number of skewed version of the GARCH models have been proposed. See Miron & Tudor (2010) for a survey of asymmetric GARCH models. In section 4.1 of this thesis the Heston Nandi GARCH (HNG) model, as introduced by Heston & Nandi (2000), is employed.

[b] A random walk is defined as a process $Y_T = \sum_{t=1}^{T} X_t$ with X_t is an i.i.d. process and $Y_0 = 0$. White noise log returns therefore correspond to the setting where the logarithmic asset value process follows a random walk process with white noise increments.

[c] Stylised facts are typical features that are empirically observed in a large portion of financial time series.

[d] Strict positivity holds if $\alpha_1 > 0$.

[e] The kurtosis $K(X)$ of a random variable X is defined as the standardised fourth moment, i.e. $K(X) = \mathbb{E}\left[\frac{(X-\mathbb{E}[X])^4}{(\mathrm{Var}(X))^2}\right]$, excess kurtosis is defined as $K(X) - K(Z) = K(X) - 3$ where $Z \sim \mathcal{N}(0,1)$.

4. Model formulation and estimation

Throughout this section we let $(\Omega, \mathcal{F}_T, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P})$ be a filtered probability space on a finite time horizon $[0, T]$ that characterises the market. Here Ω represents all available states, $(\mathcal{F}_t)_{t\in[0,T]}$ is a filtration, where $\mathcal{F}_t$ can be interpreted as the information set up to time t so that $\mathcal{F}_T$ is the information set at the end of the horizon, and $\mathbb{P}$ is the physical probability measure. It is assumed that the market is free of arbitrage. By the first fundamental theorem of asset pricing there then exists a (not necessarily unique) risk neutral probability measure[a] $\mathbb{Q}$ that is equivalent to $\mathbb{P}$. $\mathcal{F}_t$-conditional expectation under $\mathbb{P}$ and $\mathbb{Q}$ will be denoted by $\mathbb{E}_t[\cdot]$ and $\mathbb{E}^*_t[\cdot]$ respectively.

4.1. Heston Nandi GARCH Model

The Heston & Nandi (2000) GARCH (HNG) model is defined by the following assumption:

Assumption 1. The total return asset time series $(S_t)_{t\in[0,T]}$ evolves discretely with time steps Δ by the following process:

$$\begin{aligned} \ln\left(\frac{S_t}{S_{t-\Delta}}\right) &= r_{t-\Delta} + \lambda h_t + \sqrt{h_t} z_t \\ h_t &= \omega + \beta h_{t-\Delta} + \alpha(z_{t-\Delta} - \gamma\sqrt{h_{t-\Delta}})^2, \end{aligned} \tag{4.1}$$

with r_t the continuously compounded interest rate per time step, that is known at t and constant on time interval $[t, t+\Delta]$, and where z_t is a standard normal innovation term conditional on $\mathcal{F}_{t-\Delta}$. The quantity h_t is the volatility of the log return between $t-\Delta$ and t, which is predictable in the sense that it is $\mathcal{F}_{t-\Delta}$-measurable.

From (4.1) we note that the skewness parameter γ distinguishes the HNG model from the classical GARCH model by Bollerslev (1986) (GARCH(1,1) as defined in (1) is recovered for $\gamma = 0$). As the $\mathcal{F}_{t-\Delta}$-conditional covariance between $\ln S_t$ and $h_{t+\Delta}$ is equal to $-2\alpha\gamma h_t$ (see appendix D.1), a positive γ

[a] A probability measure $\mathbb{Q}$ is called risk-neutral if the discounted value process of any self-financing portfolio has the martingale property under $\mathbb{Q}$. That is, for every self-financing portfolio, it holds that

$$e^{-r(T-t)}\mathbb{E}_Q[V_T|\mathcal{F}_t] = V_t.$$

where r denotes the risk free rate.

enables the model to capture the skewness, or leverage effect (Black 1976), that is observed both in option prices as well as in historical realisations.

Hereafter the parameter vector, as estimated under physical measure $\mathbb{P}$, will be denoted by

$$\theta := (\omega, \alpha, \beta, \gamma, \lambda). \tag{4.2}$$

4.1.1. Equilibrium volatility

It is seen in Heston & Nandi (2000) that the persistence of the HNG-process is $\gamma + \alpha\gamma^2$, so that the process is stationary if $\gamma + \alpha\gamma^2 < 1$ (see Francq & Zakoïan (2010) for details). To find the long term equilibrium volatility, in the second equality of (4.1) set $h_t = h_{t+1} = h_{eq}$ and take expectation to obtain (using that $z_t \sim \mathcal{N}(0,1)$):

$$\begin{aligned} h_{eq} &= \mathbb{E}[\omega + \beta h_{eq} + \alpha(z_{t-\Delta} - \gamma\sqrt{h_{eq}})^2] \\ &= \omega + \beta h_{eq} + \alpha\gamma^2 h_{eq} + \mathbb{E}[\alpha z_{t-\Delta}^2 - 2\alpha\gamma z_{t-\Delta}\sqrt{h_{eq}}] \\ &= \omega + \beta h_{eq} + \alpha\gamma^2 h_{eq} + \alpha \end{aligned}$$

so that

$$h_{eq} = \frac{\omega + \alpha}{1 - \alpha\gamma^2 - \beta}. \tag{4.3}$$

Observe from (4.3) that the stationarity condition $\beta + \alpha\gamma^2 < 1$, together with $\omega + \alpha > 0$, is sufficient to ensure positive equilibrium volatility.

4.2. Pricing kernel

In order to be able to price options with the HNG-model, we need to define a pricing kernel.

Definition 2. A pricing kernel M_t is a strictly positive $(\mathcal{F}_t)_{t\in[0,T]}$-adapted random variable such that the value at t of any pay-off π at future time $T \geq t$ is equal to $\mathbb{E}_t^*[\pi] = \mathbb{E}_t[M_t \cdot \pi]$[b].

Christoffersen et al. (2013) generalise the model as defined by Heston & Nandi (2000) by defining the following variance-dependent pricing kernel:

$$M_t = M_0 \left(\frac{S_t}{S_0}\right)^{\psi} \exp\left(\delta t + \eta \sum_{s=0}^{t} h_s + \xi(h_{t+\Delta} - h_0)\right) \tag{4.4}$$

for some[c] $(\xi, \delta, \eta, \psi) \in \mathbb{R}^4$, and with S_t the value of the underlying asset, h_t the volatility at time t as before. Christoffersen et al. (2013) prove the following proposition (see appendix D.6 for details of this proof):

[b] See Leisen (2017) for an overview of the basic properties of the pricing kernel.

[c] Here δ, η, ψ are just algebraic re-orderings of $\mathbb{P}$-parameters θ and ξ, see appendix D.6.

Proposition 1. *With pricing kernel* (4.4) *the risk-neutral stock process corresponding to Heston Nandi GARCH process* (4.1) *is the GARCH process:*

$$\begin{aligned} \ln\left(\frac{S_t}{S_{t-\Delta}}\right) &= r_{t-\Delta} - \frac{1}{2}h_t^* + \sqrt{h_t^*}z_t^* \\ h_t^* &= \omega^* + \beta h_{t-1}^* + \alpha^*\left(z_{t-1}^* - \gamma^*\sqrt{h_{t-1}^*}\right)^2 \end{aligned} \tag{4.5}$$

where z_t^* *is standard normal under* $\mathbb{Q}$ *and*

$$\begin{aligned} h_t^* &= h_t/(1-2\alpha\xi) \\ \omega^* &= \omega/(1-2\alpha\xi) \\ \alpha^* &= \alpha/(1-2\alpha\xi)^2 \\ \gamma^* &= (\lambda+\gamma)(1-2\alpha\xi) + \frac{1}{2}. \end{aligned} \tag{4.6}$$

It is well documented, see for example Carr & Wu (2009), that in the equity option market, a negative risk premium exists (that is, 'Variance buyers are willing to accept a negative average excess return to hedge away upward movements in stock market volatility' (Carr & Wu 2009)). This corresponds to a strictly positive ξ, so that the variance of the log return process is higher under $\mathbb{Q}$ than it is under $\mathbb{P}$.

The original Heston & Nandi (2000) model corresponds to the special case of this generalised HNG-model with $\xi = 0$. In this version of the model volatility under $\mathbb{Q}$ is identical to $\mathbb{P}$-volatility, and (4.1) holds under $\mathbb{Q}$ with $\lambda^* = -\frac{1}{2}$ and $\gamma^* = \gamma + \lambda + \frac{1}{2}$. It is seen in Heston & Nandi (2000) that this special case corresponds to the assumption that the Black-Scholes pricing formula (4.15) holds on a single time-step. Hereafter the vector of $\mathbb{Q}$-parameters[d] is denoted by

$$\theta^* := (\omega^*, \alpha^*, \beta, \gamma^*). \tag{4.7}$$

It is noted from (4.6) that the risk neutral parameter space is uniquely defined as a function from physical parameter vector (4.2) and ξ. That is, there exists risk-neutral parameter vector

$$\theta^* = (\omega^*, \alpha^*, \beta, \gamma^*) = F(\theta, \xi) = F(\omega, \alpha, \beta, \gamma, \lambda, \xi)$$

where $F : \mathbb{R}^6 \to \mathbb{R}^4$ is defined by the last three equalities of (4.6). It is remarked that for the special case $\xi = 0$, the correspondence between the $\mathbb{P}$-process and the $\mathbb{Q}$-process is 1-to-1 (setting $\lambda^* = -\frac{1}{2}$ and $\gamma^* = \gamma + \lambda + \frac{1}{2}$, as follows from (4.6)), and that in particular no model calibration under $\mathbb{Q}$ is required to obtain a risk-neutral version of the process.

Now that a risk-neutral formulation of the HNG-model is found, an option pricing formula can be derived.

[d] Note that price of risk parameter λ does not play a role in the risk neutral setting.

4.3. Option pricing under the HNG-model

4.3.1. Option pricing formula

Heston & Nandi (2000) show that the $\mathcal{F}_t$-conditional moment generating function of the logarithm of the asset price under $\mathbb{P}$, that is

$$f(\phi) = \mathbb{E}_t[(S_T)^\phi], \tag{4.8}$$

takes a log linear form:

$$f(\phi) = (S_t)^\phi \exp\left(A(t,T;\phi) + B(t,T;\phi)h_{t+\Delta}\right) \tag{4.9}$$

where $A(t,T;\phi)$ and $B(t,T;\phi)$ can be recursively derived from the terminal conditions $A(T,T;\phi) = B(T,T;\phi) = 0$. Furthermore, the following closed form option pricing formula is derived:

$$\begin{aligned}\mathbb{E}_t[(S_T - K)^+] = f(1)\left(\frac{1}{2} + \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left[\frac{K^{-i\phi}f(i\phi+1)}{i\phi f(1)}\right]\mathrm{d}\phi\right) \\ - K\left(\frac{1}{2} + \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left[\frac{K^{-i\phi}f(i\phi)}{i\phi}\right]\mathrm{d}\phi\right)\end{aligned} \tag{4.10}$$

with $f(\cdot)$ the conditional moment generating function of the log asset price (4.8), and where Re$[arg]$ denotes the real part of arg. For a European call option with strike price K we then have

$$\begin{aligned}C &= e^{-r_t(T-t)}\mathbb{E}_t^*[(S_T - K)^+] \\ &= \frac{1}{2}S_t + \frac{e^{-r_t(T-t)}}{\pi}\int_0^\infty \mathrm{Re}\left[\frac{K^{-i\phi}f^*(i\phi+1)}{i\phi}\right]\mathrm{d}\phi \\ &\quad - Ke^{-r_t(T-t)}\left(\frac{1}{2} + \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left[\frac{K^{-i\phi}f^*(i\phi)}{i\phi}\right]\mathrm{d}\phi\right)\end{aligned} \tag{4.11}$$

where $f^*(\phi)$ denotes the $\mathcal{F}_t$-conditional moment generating function under $\mathbb{Q}$. The details of the derivation of the recursive expression for (4.8) and of the option pricing formulas (4.10) and (4.11) are presented in appendix D.

Call option pricing formula (4.11) takes volatility $h_{t+\Delta}$ as input (implicitly through (4.9)). Throughout this thesis, the equilibrium volatility will be used for the computation of option prices. From the second equality in (4.5), analogue to (4.3), h^*_{eq} is found:

$$h^*_{eq} = \frac{\omega^* + \alpha^*}{1 - \alpha^*\gamma^{*2} - \beta}. \tag{4.12}$$

4.3.2. Dividend correction

The option pricing formula (4.11) assumes no dividends, whereas dividend is paid to shareholders by at least some of the assets that make up the S&P 500 index. The ICE data (see section 2) provides estimates of the (annualised)

value of the dividend that is paid out during the maturity of the option. Following the approach of among others Christoffersen et al. (2013) and Heston & Nandi (2000), a dividend adjusted index value of the underlying stock is used for the computation of option prices. This is found by deducting the present day value of the expected dividend pay-out during the maturity of the option from the current underlying index value. Assuming continuously compounded dividend pay-out at rate q until maturity time T this amounts to correcting for the dividend drift of the asset by using

$$\tilde{S}_t = \frac{S_t}{e^{qT}}$$

as input for the computation of the option price. See appendix D.2 for an intuitive verification of this dividend correction.

4.4. Historic calibration

The physical parameter vector (4.2) is estimated on a time series of historical S&P 500 log returns $R := (R_t)_{1 \leq \tau \leq t}$ and a time series of historical interest rates $r = (r_\tau)_{1 \leq \tau \leq t}$ (see section 2). The total log likelihood of (4.1) is equal to

$$LL_{\mathbb{P}}(\theta, R, r) = \sum_{\tau=1}^{t} \left[-\frac{1}{2} \ln(2\pi h_\tau) - \frac{(R_\tau - r_\tau - \lambda h_\tau)^2}{2h_\tau} \right], \tag{4.13}$$

where current time t is the length of time series R and r. Log likelihood (4.13) is found by isolating z_τ and using the standard normal density function. Given historical data set (R, r), we then numerically optimise to find:

$$\hat{\theta}_{hist}(R, r) := \arg\max_{\theta} LL_{\mathbb{P}}(\theta, R, r). \tag{4.14}$$

The Nelder-Mead algorithm is used to find (4.14). The first column in table B.1 shows the estimated parameters. See figure 4.1 for a plot of a simulation with these parameters.

4.5. Implied volatilities

As defined by Black & Scholes (1973) and generalised to account for dividend by Merton (1973), the Black-Scholes pricing formula for European call option with strike K, time to maturity $T - t$, risk free rate r, current stock price S_t, annual dividend rate q is:

$$C_{BS}(\sigma, T-t, K, r) = \mathcal{N}(d_1) S_t e^{-q(T-t)} - \mathcal{N}(d_2) K e^{-r(T-t)} \tag{4.15}$$

with $\mathcal{N}$ the cumulative standard normal distribution and

$$d_1 = \frac{\ln(S_t/K) + (r - q + \sigma^2/2)(T-t)}{\sigma\sqrt{t-t}}, \quad d_2 = d_1 - \sigma\sqrt{T-t}.$$

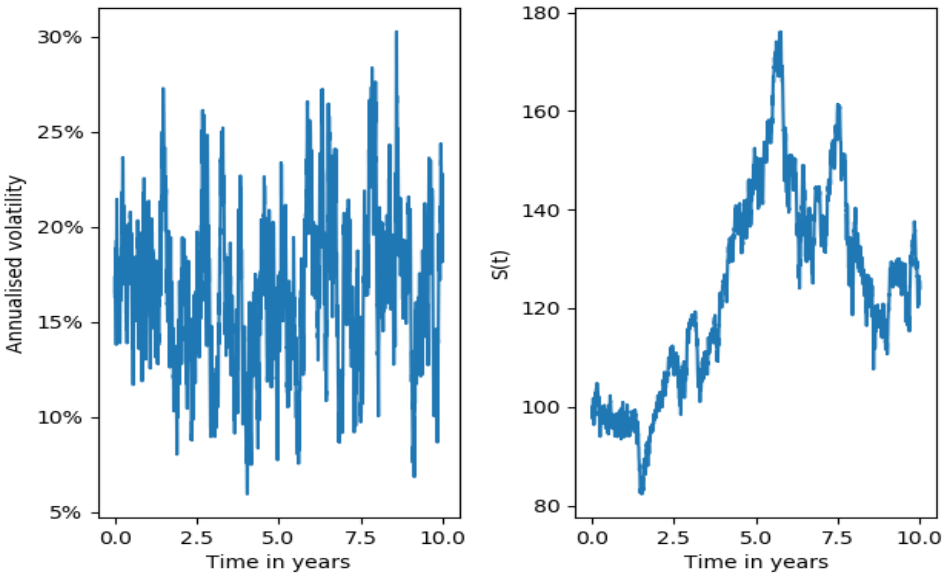

Figure 4.1. HNG model simulation
Simulation of annualised volatility and stock value in HNG process on ten year horizon, with parameters $\hat{\theta}_{hist}(R) = \{\omega : 1.2 \cdot 10^{-11}, \alpha : 4.0 \cdot 10^{-6}, \beta : 0.9, \gamma : 189.2, \lambda : 2.2\}$ with R S&P 500 data ranging from 1990-01-01 to 2016-12-31.

Per convention the implied volatility (IV) of a certain option price is equal to the input volatility for the Black-Scholes option pricing formula that would give this option price. That is, for HNG call price C_{HNG} with relative moneyness m (i.e. $K = m \cdot S_t$), we define:

$$IV_m^{HNG} := \{\sigma \in \mathbb{R}_+ : C_{BS}(\sigma, T-t, K = m \cdot S_t, r) = C_{HNG}\} \tag{4.16}$$

where C_{BS} is the call value under Black-Scholes formula (4.15). As evaluation of goodness of fit, the implied volatility skew[e] of HNG option prices is compared to market implied volatility skews, as it is of interest whether the HNG model reaches the right level of market implied volatility, and is able to capture the pronounced skewness that is observed in the market.

4.6. Recalibration to market implied volatility skew

A second calibration procedure is performed by fitting the model with respect to market implied volatilities[f], i.e. under $\mathbb{Q}$. As was announced in the introduction, this will firstly be done by freely fitting θ^* under $\mathbb{Q}$ ('ad-hoc method', see section 4.6.1), which will serve as benchmark. Secondly a sequential method

[e] Implied volatilities plotted against maturity

[f] This is an indirect calibration approach, as opposed to the approach followed by Heston & Nandi (2000) and Christoffersen et al. (2013), who both minimise errors with respect to option prices 'directly'. Here we chose to calibrate with respect to IV's, as the ICE data consists of IV's and by this approach any scaling issues are avoided.

is applied (see section 4.6.2). We use N different market data points. For each data point implied volatilities are available for maturities between 1 year and 10 years, and with moneyness factor $(x_i)_{1\leq i\leq I}$ ranging from $x_1 = 0.5$ to $x_I = 1.5$ (i.e. strike price $K = x_i \cdot S_t$, so that e.g. call options are 50% in the money for moneyness factor $x_1 = 0.5$).

4.6.1. Ad-hoc method

Given parameter vector θ^*, implied volatilities of HNG call option are computed for some subset of the available moneyness range $\{x_i : 1 \leq i_m \leq i \leq i_M \leq I\}$. For the moneyness range, a distinction will be made between *near the money* (NTM) calibration, where $x_{i_m} = 0.9 \leq x_i \leq 1.1 = x_{i_M}$ and *away from the money* (AFTM) calibration, where $x_{i_m} = 0.5 \leq x_i \leq 1.5 = x_{i_M}$. Given market dataset $IV^{market} = \{IV_n^{market}\}_{1\leq n\leq N}$ we define:

$$\varepsilon_{x_i}(\theta^*, IV_n^{market}) := IV_{x_i,n}^{market} - IV_{x_i}^{HNG}(\theta^*). \tag{4.17}$$

Following Christoffersen et al. (2013), the criterion function[g] is derived by assuming that the square of these errors are i.i.d. normal with expectation zero and some variance σ_ε^2. Writing $(\varepsilon_{x_i}(\theta^*, IV_n^{market}))^2$ as $\varepsilon_{x_i}^2(\theta^*, IV_n^{market})$, we get the total log likelihood with respect to market data set n:

$$LL_{\mathbb{Q}}(\theta^*, IV_n^{market}, \sigma_\varepsilon^2) = \sum_{i=i_m}^{i_M} -\frac{1}{2}\ln(2\pi\sigma_\varepsilon^2) - \frac{1}{2}\varepsilon_{x_i,n}^2(\theta^*, IV_n^{market})/\sigma_\varepsilon^2 \tag{4.18}$$

and we find ad-hoc risk neutral parameters $\hat{\theta}_{ad-hoc}^*$ by setting:

$$\hat{\theta}_{ad-hoc}^*(IV^{market}) := \arg\max_{\theta^*} \sum_{n=1}^{N} LL_{\mathbb{Q}}(\theta^*, IV_n^{market}, \hat{\sigma}_\varepsilon^2(\theta^*, IV_n^{market})) \tag{4.19}$$

with

$$\hat{\sigma}_\varepsilon^2(\theta^*, IV_n^{market}) := \frac{1}{N(i_M - i_m + 1)} \sum_{n=1}^{N} \sum_{i=i_m}^{i_M} \varepsilon_{x_i}^2(\theta^*, IV_n^{market}).$$

4.6.2. Sequential parameter estimation

In the sequential parameter estimation approach, we define the $\mathbb{Q}$-parameter vector as

$$\hat{\theta}_{seq}^* = F(\hat{\theta}_{hist}(R, r), \hat{\xi}(IV^{market}))$$

[g] Alternatively, one could follow Heston & Nandi (2000) by minimising the sum of squared errors (4.17), to obtain

$$\hat{\theta}_{IV,error}(IV^{market}) := \arg\min_{\theta^*} \sum_{n=1}^{N} \sum_{i=i_m}^{i_M} \varepsilon_{x_i}^2(\theta^*, IV_n^{market}).$$

Calibration results turn out to be almost identical to results when using the Christoffersen et al. (2013) log likelihood criterion function (available on request).

where F is the mapping as introduced in section 4.2, $\hat{\theta}_{hist}(R, r)$ is the $\mathbb{P}$ parameter vector (4.14), and $\hat{\xi}$ is found by setting (dropping dependencies for $\hat{\theta}_{hist}$ and $\hat{\sigma}_{\varepsilon}^2$):

$$\hat{\xi}(IV^{market}) = \arg\max_{\xi} \sum_{n=1}^{N} LL_{\mathbb{Q}}(F(\hat{\theta}_{hist}, \xi), IV_n^{market}, \hat{\sigma}_{\varepsilon}^2)$$

with $LL_{\mathbb{Q}}(\theta^*, IV_n^{market})$ defined in (4.18).

4.7. Recalibration to OFS

Next, the $\mathbb{P}$-parameters are recalibrated on forward looking information instead of historical realisations. For this purpose the economic scenarios from the Ortec Finance Scenario set (OFS, see section 2.1) are used. We use a set of 2000 scenarios for a ten years horizon for total return US equity, and a corresponding set of 2000 scenarios for the 3-month US swap rate, both from the December 2018 scenario set, denoted R and r respectively. As these scenarios are monthly, this procedure will only be performed for formulation of the HNG-model with monthly or yearly frequency (see section 4.8). We define newly calibrated real world parameters:

$$\hat{\theta}_{scen}(R, r) := \arg\max_{\theta} \sum_{n=1}^{2000} LL_{\mathbb{P}}(\theta, R_n, r_n) \tag{4.20}$$

where $LL_{\mathbb{P}}$ was defined in (4.13), R_n is the n-th scenario for US equity total return, and r_n is the n-th scenario for the 3-month US swap rate.

Subsequently the sequential fit method from section 4.6.2 is used to find an estimate of the parameter vector for what we will call the scenario-sequential model:

$$\hat{\theta}^*_{scen} = F(\hat{\theta}_{scen}(R, r), \tilde{\xi}(IV^{market})) \tag{4.21}$$

with (again dropping dependencies for $\hat{\theta}_{scen}$ and $\hat{\sigma}_{\varepsilon}^2$)

$$\hat{\xi}(IV^{market}) = \arg\max_{\xi} \sum_{n=1}^{N} LL_{\mathbb{Q}}(F(\hat{\theta}_{scen}, \xi), IV_n^{market}, \hat{\sigma}_{\varepsilon}^2).$$

4.8. Model frequency

Whereas GARCH models are typically formulated in daily time-steps, on a very long maturity horizon it is of interest to evaluate whether reformulation of the (generalised) HNG-model in monthly and yearly time-steps might enable the model to better grasp long maturity effects. Results are given in section 5.4. We will use a rolling window approach for the resampling of historical data and of the economic scenarios.

5. Results

5.1. Historically calibrated model

The parameter vector under $\mathbb{P}$, (4.14), is estimated on a time series of daily S&P 500 total returns between 1990 and 2016. See table B.1 for the estimates.

Figure 5.1 shows the IV-skew of the historically calibrated HNG model, compared with market IV-skew on a ten year horizon. Despite the trend being similar (decreasing and convex) it is clear that the historically calibrated HNG does not come near capturing a skew as large as is seen in the market, and that it is significantly lower for every moneyness.

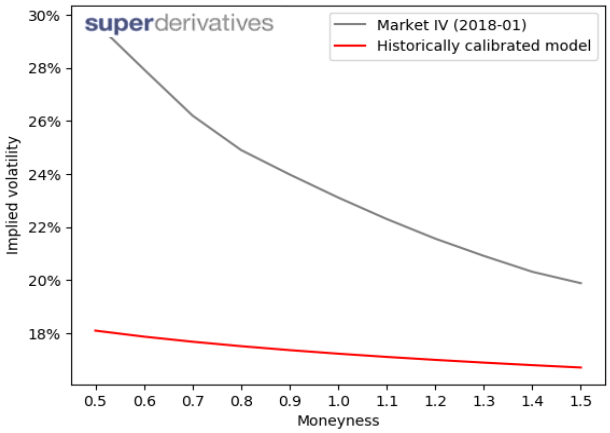

Figure 5.1. IV skew under $\mathbb{P}$

Implied volatility from HNG call price with maturity $T = 10$, with parameter vector $\hat{\theta}^*_{hist}$ (risk neutral version, transformed to $\mathbb{Q}$ under standard HNG-model, i.e. with $\xi = 0$, of parameter vector $\hat{\theta}_{hist}$ that was calibrated to S&P 500 data between 1990 and 02/2020), plotted with market IV skew (ICE data from January 2018). On the x-axis is the relative moneyness m ($K = m \cdot S_t$ with K strike price and S_t current asset value). Source market data: Superderivatives.

5.2. The ad-hoc benchmark

The (daily) ad-hoc model parameter vector (4.19) is estimated by calibrating against market implied volatilities between 2017-07 and 2018-12. The estimate can be found in table B.1. Option prices and corresponding implied volatilities are computed for relative moneyness between 0.5 and 1.5. As is seen in figure 5.2 this model is able to give a good approximation of the market IV-skew.

Notably the modelled implied volatility and market implied volatility match better as maturity increases.

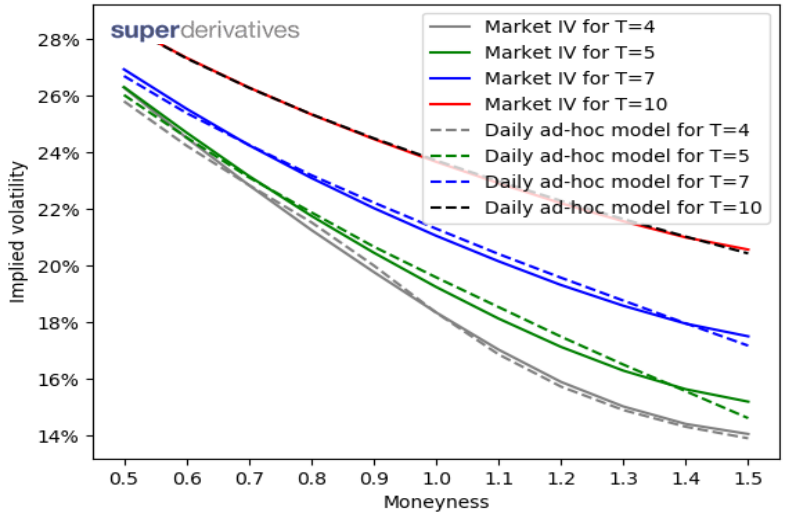

Figure 5.2. Ad-hoc model
IV-skews of daily ad-hoc HNG-model plotted with in sample market IV. The model was fitted on IV data between 2017-07 and 2018-12. The plotted market IV are average IV's over this in-sample period. On the x-axis is the relative moneyness m ($K = (1+m)S_t$ with K strike price and S_t current asset value). Source market data: Superderivatives.

5.3. Sequential model

Parameter estimates for the ad-hoc model and the sequential model are displayed for maturity $T = 10$ years in table B.1. We note that the risk-neutral equilibrium volatility under the ad-hoc model is substantially higher than the equilibrium volatility under $\mathbb{P}$. The sequential estimation partially bridges this gap, but under unrestricted $\mathbb{Q}$-calibration of the ad-hoc model, the risk neutral equilibrium volatility remains significantly higher. Furthermore we note that the risk neutral skewness parameter γ^* is much larger in the ad-hoc model.

The total log likelihood (LL)[a] is significantly larger for the ad-hoc model than it is for the sequential model, reflecting a much better fit. This is confirmed by a plot of the IV-skews for $T = 10$ in figure 5.3. Comparing with the fit of the strictly $\mathbb{P}$-calibrated model from figure 5.1 we observe that the $\mathbb{P}/\mathbb{Q}$ mapping by $\hat{\xi}$ was able to overcome the difference in height, but the sequential model comes up well short in approximating the market skewness. This is reflected in the considerable difference in root mean squared error (RMSE) for away from the money options (0.6147 for ad-hoc model versus 2.179 for the sequential model), whereas for near the money options the difference in RMSE is much smaller (0.6409 and 0.8036 respectively).

[a]Not to be compared with log likelihood under $\mathbb{P}$-calibration in the first column.

The same $\mathbb{Q}$-calibrations were performed against options with maturity varying between 2 and 10 years[b]. Figure 5.4 confirms similar findings for these maturities, with away from the money errors significantly higher for the sequential method.

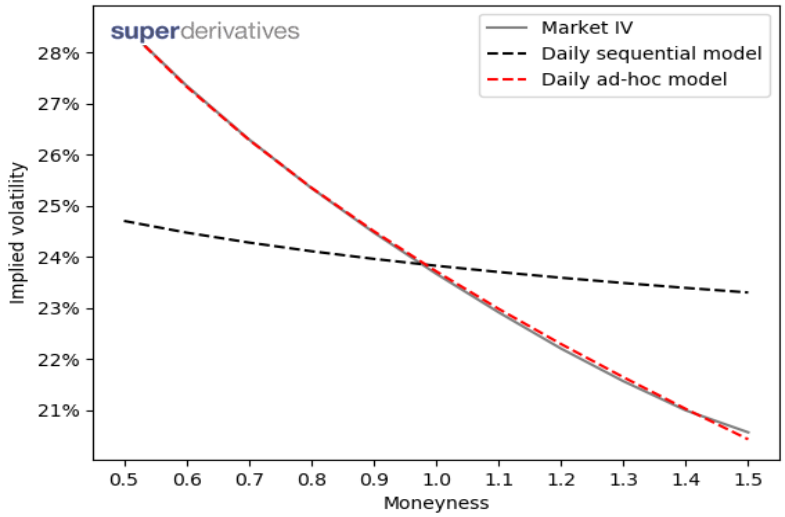

Figure 5.3. Ad-hoc and sequential model
IV-skews of daily ad-hoc HNG-model ($\hat{\theta}^{*AFTM}_{ad-hoc}$) and daily sequential HNG-model ($\hat{\theta}^{*AFTM}_{seq}$) plotted with in sample mean of market IV. Source market data: Superderivatives.

5.3.1. Side note regarding the dataset

Having computed estimates $\hat{\xi}$ for a range of maturities, an intermediate remark is made: In Christoffersen et al. (2013) a negative volatility premium ξ, with corresponding ξ-factor $(1-2\alpha\xi)^{-1}$ around 1.3 was estimated. In this thesis for shorter maturity a positive volatility premium (ξ-factor $(1-2\hat{\alpha}\hat{\xi})^{-1} < 1$, or equivalently $\hat{\xi} < 0$) was found, and the ξ-factor was estimated close to 1 for $T = 3$ (see the left plot in figure 5.11). This reflects the atypical nature of the data set that was used. In the literature typically a diverse and large array set of short maturity option data is used, discarding illiquid products. In this research a data set was used where for each month only a single implied volatility point was available per moneyness/maturity, for a wide arrange of maturities and moneyness. This data set contained a relatively small number of near the money options, and contained very far out of/in the money options (up to $\pm 50\%$). A large skewness was observed in this market data (figure 5.2), in particular for shorter maturity options, where far in/out of the money options represent extremely illiquid products. In the HNG-model, the γ-parameter was implemented to cover this skewness, and from (4.6) it

[b]Results available on request.

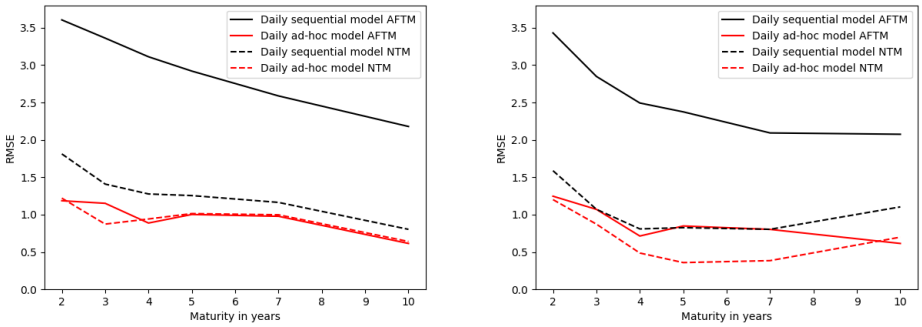

Figure 5.4. In sample (left) and out of sample (right) RMSE for daily models against T

The RMSE of model IV with respect to market IV is plotted against maturity in years T for the sequential and ad-hoc (daily) HNG-model. The in sample period ranges from 2017-07 to 2018-12, and the out of sample period from 2019-01 to 2020-02.

is observed that (for positive α, λ, γ) a larger ξ leads to smaller risk neutral skewness γ^*, explaining why $\hat{\xi}$ tends to negativity for smaller maturity.

5.4. Monthly and yearly reformulation

The model is reformulated to monthly and yearly frequencies[c]. For the historical calibration a rolling window approach is applied. The estimated monthly and yearly parameters (where again $\mathbb{Q}$-calibration is done for $T = 10$) are found in tables B.2 en B.3 (the tables were confined to $\mathbb{Q}$-calibration against away from the money options). This table also presents results for calibration against economical scenarios (see section 5.6).

The yearly model turns out to consistently outperform the daily model, both in the case of the sequential and in the case of the ad-hoc model, while the monthly model consistently performs worst. It is seen in figure 5.5 that the yearly model obtains more skewness than the other frequencies. From table B.3 it is observed that $\hat{\gamma}$ is small in the yearly $\mathbb{P}$-model, which shows that in the HNG-model little skewness is detected from the history of annual stock returns. However, γ^* is relatively large, resulting in a more skewed model under $\mathbb{Q}$, compared to its daily counterpart.

5.5. Varying start year in historical $\mathbb{P}$-calibration

For the sequential model, the first calibration step, the historical $\mathbb{P}$-estimation, was thus far been done with respect to S&P 500 log returns between 1990 and 2016. This start year was varied in order to get a more robust comparison of model performance for different frequencies, and to find an optimal starting

[c]More experiments with varying frequencies (3/6-monthly, 2/5/10-yearly) were conducted. Error margins of these models are comparable, but consistently inferior to those of the yearly model. Details are available on request.

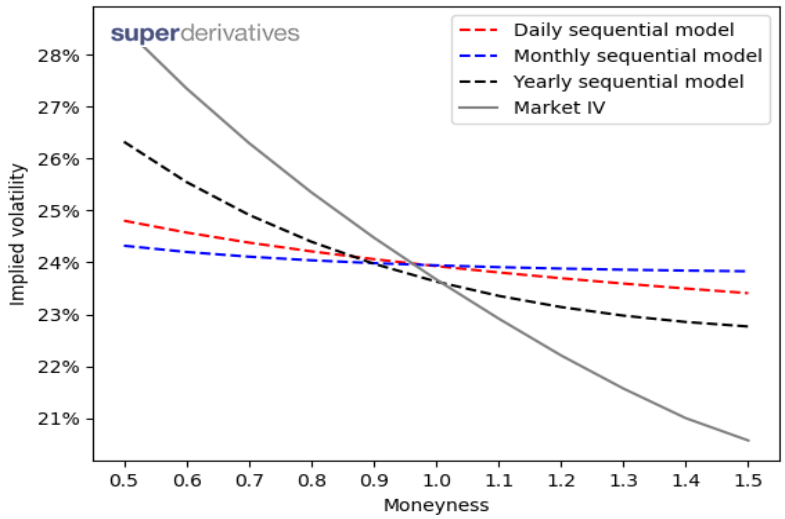

Figure 5.5. IV-skews for different frequencies of sequential model
Mean of in-sample market IV-skews is displayed along with IV-skew from daily/monthly/yearly version of HNG sequential model for maturity of ten years. Source market data: Superderivatives.

year. An error comparison for the daily and yearly model[d] with starting years between 1950 and 2005 is seen in figure 5.6.

A number of observations are made: First we note that in particular for the yearly model and regarding the out of sample errors, the starting year has a significant influence. An optimum seems to be found for the yearly model by using start year between 1976 and 1991, i.e. using a historical calibration window that is between 40 and 25 years long, while the daily model reaches a minimum in RMSE with a 16 year data range. On the optimal interval for the yearly model, the daily model is comfortably outperformed, both in terms of near the money and away from the money errors.

5.6. The scenario-sequential model

As the economical scenarios from the Ortec Finance Scenario set (OFS) are formulated monthly, we will only perform recalibration to economical scenarios for the monthly and yearly frequency formulations of the model. Both version of (4.20) ($\hat{\theta}^M_{scen}$ and $\hat{\theta}^Y_{scen}$ respectively) for $T = 10$ are found in tables B.2 and B.3. It is observed that, consistent with the regular sequential model, error margins (Log Likelihood, RMSE etcetera) for the yearly version of the model were better than for the monthly model.[e]

[d] The monthly model was consistently outperformed, and is left out of the plot for clarity (available on request).

[e] This holds not only for the in sample errors as displayed in tables B.2 and B.3, but also for out of sample errors.

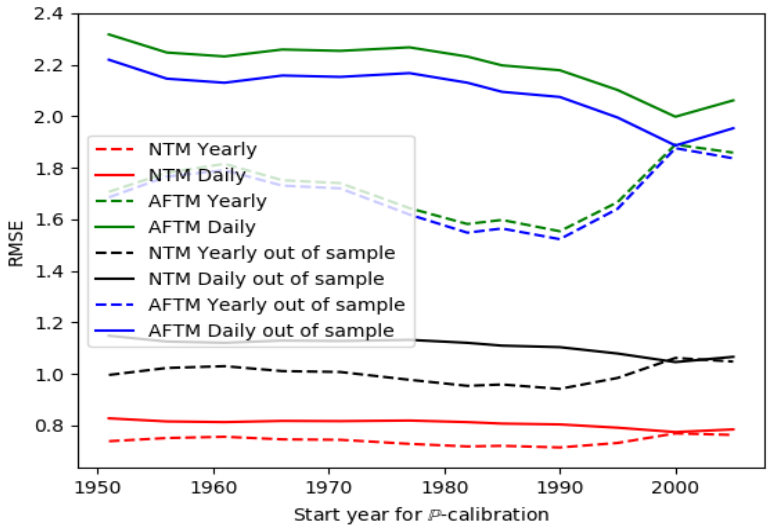

Figure 5.6. RMSE plotted against start year for historical calibration
In sample/out of sample RMSE's of the HNG-model IV with respect to market IV is plotted for away from the money/near the money options.

The error evaluation of figure 5.4 is repeated for the yearly sequential model and the yearly scenario-sequential model in figures 5.7 - 5.10. The out of sample analogues are found in C.1 - C.4. The yearly model consistently outperforms the daily model for maturity of at least three years. It is furthermore observed that both the in and the out of sample errors of the scenario-sequential model consistently improve with respect to those of the regular yearly sequential model. While in-sample the ad-hoc model holds a small bias and the sequential model is virtually unbiased, an out-of-sample bias is present for the sequential models and smaller for the ad-hoc models.

Results of calibration against longer maturity horizon (20/30 years) scenarios, with comparison to IV-skews of the Risk-Neutral Heston (1993) model is given in appendix A.

5.7. Evolution of $\mathbb{Q}/\mathbb{P}$ mapping

The $\mathbb{P}/\mathbb{Q}$ mapping as defined by (4.4) is estimated by fitting $\hat{\xi}$ under $\mathbb{Q}$, and is characterised (see (4.6)) by the mapping factor $(1-2\hat{\alpha}\hat{\xi})^{-1}$. It is evaluated how the estimate of this ξ-factor evolves for increasing maturity. Three plots of the fitted $\hat{\xi}$ and the corresponding factors are plotted in figure 5.11. It is seen that the ξ-factor evolves nearly linearly, while the $\hat{\xi}$ is somewhat concave. This pattern was not only seen across the three model variants plotted in figure 5.11 (the regular sequential model for three different frequencies, fitted against away from the money options), but was very similar for the scenario-sequential model in both monthly and yearly frequency, and also for either of

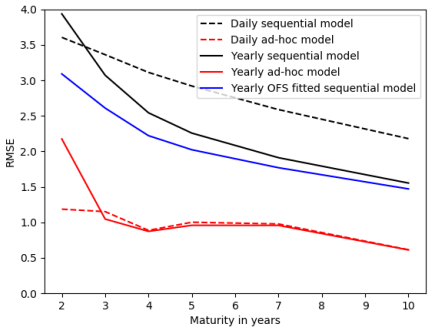

Figure 5.7. In sample away from the money RMSE

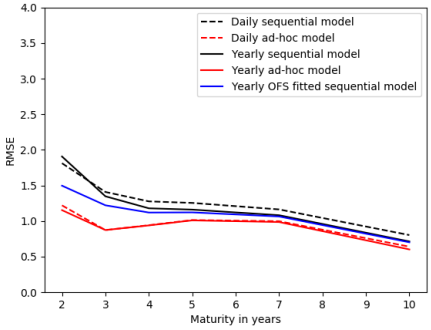

Figure 5.8. Near the money RMSE

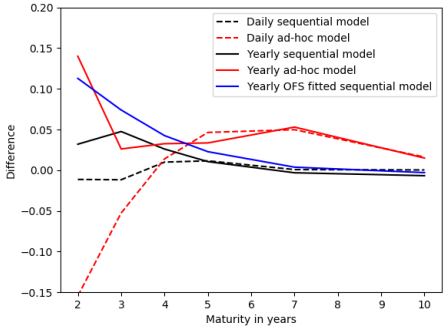

Figure 5.9. Away from the money bias

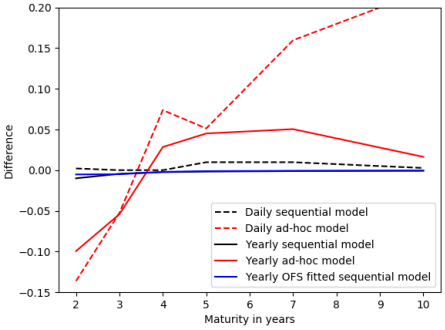

Figure 5.10. Near the money bias

The top two figures show in sample RMSE of model IV with respect to market IV is plotted against maturity in years T for the sequential and ad-hoc HNG-models with varying frequency. The bottom two figures show the mean model bias, i.e. $IV^{Market} - IV^{HNG}$. The in sample period ranges from 2017-07 to 2018-12.

the models with $\mathbb{Q}$-calibration confined to near the money options.

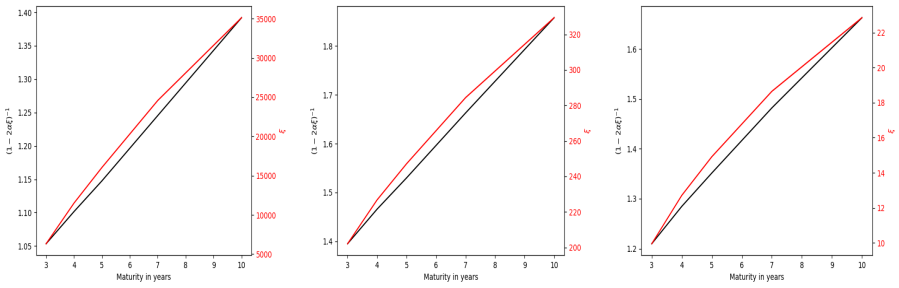

Figure 5.11. Three plots of $\hat{\xi}$ (with scale on right y-axis) and $\hat{\xi}$-factor (scale on left y-axis) against maturity T

Three different plots of estimates of mapping parameter ξ, with the corresponding estimate of the factor $(1-2\alpha\xi)^{-1}$ are plotted. From left to right $\hat{\xi}$ if plotted for the daily, monthly and yearly regular sequential HNG-model.

6. Discussion

In this thesis, the Heston & Nandi (2000) GARCH model was implemented in the context of long maturity (2-10 years), and the efficiency of the variance dependent pricing kernel by Christoffersen et al. (2013) was empirically scrutinised, as well as its consistency along increasing maturity. The general finding that a $\mathbb{Q}$-calibrated HNG-model is able to match market option prices well, was confirmed for the long maturity context.

The available option data showed a severe skewness, with the level of the implied volatility surface increasing with T. The $\mathbb{P}$-calibrated HNG-model considerably underestimates the skewness in the market, and strongly undervalues the volatility under $\mathbb{Q}$. This first observation indicates that the implied volatility skew as present in the market is only partially explained by the leverage effect in asset returns, as is confirmed by Kozhan et al. (2013).

Whereas the sequential model (calibrated under $\mathbb{P}$ and transformed to $\mathbb{Q}$ with a variance premium parameter that is calibrated under $\mathbb{Q}$) comes up well short in reproducing this volatility smirk, it is found that when restricting to near the money options, error margins of the sequential model are remarkably close to the error margins of the ad-hoc model, both in and out of sample. Hereby it should be taken into consideration that the sequential model only has a single degree of freedom with respect to option prices, while the ad-hoc model has four.

After initially stating the GARCH model in the classical daily frequency, the different varieties of the model were reformulated in monthly and yearly frequency, using a rolling window approach. It was seen that on a long maturity horizon, the daily models were consistently outperformed by yearly models.

Replacing the historical calibration step by fitting against the economical scenarios of the Ortec Finance Scenario results in a model that (confined to $\mathbb{P}$-calibration, that is with $\xi = 0$) gives a more strongly skewed IV-pattern. Following up this recalibration under $\mathbb{P}$ by fitting $\hat{\xi}$ under $\mathbb{Q}$, the risk-neutral skewness is higher than $\mathbb{Q}$-skewness under the yearly regular sequential model, but still significantly flatter than the ad-hoc model.

An attractive feature of the generalised HNG-model is that it encompasses a one-dimensional characterisation of the relation between the measures $\mathbb{P}$ and $\mathbb{Q}$, in the form of the ξ-factor $(1 - 2\hat{\alpha}\hat{\xi})^{-1}$. It is found that this ξ-factor increases nearly linearly with respect to the maturity. This reflects positive correspondence between risk-aversive behaviour and maturity, with the variance premium becoming increasingly negative on a longer horizon. The latter finding is consistent with findings by amongst others Ait-Sahalia et al. (2012) and Chu, Shin-herng; Freund (1996).

The near-linearity of increase in the ξ-factor with respect to maturity is

consistent across a wide range of sequential models, with varying frequencies and $\mathbb{P}$-estimation method. Possibly one might want to extrapolate this near-linear increase to a maturity horizon longer than ten years. If one would have full confidence in the dataset[a], the diversity of the models that show almost exactly the same pattern of the mapping factor with respect to the maturity, provides fundament for extrapolation of the ξ-factor. Extrapolation of the ξ-factor, resulting in extremely large implied volatilities on a 20/30 year horizon, was applied in appendix A, together with a comparison to the Ortec Finance Risk-Neutral model (a version of the Heston (1993) model).

The atypical nature of the dataset, as outlined in 5.3.1, suited the purpose of scrutinising the ability of a generalised HNG-model to approximate a suitable $\mathbb{P}/\mathbb{Q}$-mapping in a long maturity context, where far away from the money options become more viable. However, this made it impossible to replicate typical parameter estimations, as for example found in Christoffersen et al. (2013), on shorter maturity.

6.1. Suggestions for further research

In Van Dijk et al. (2018) the benefits of an updating regime of $\mathbb{Q}$-parameters that depends on asset paths under $\mathbb{P}$ are demonstrated. A similar approach in a discrete time setting on longer maturity could prove a viable extension to the approach of this thesis.

As an alternative to the time series approach that was taken here, a similar experiment with the continuous time version of pricing kernel (4.4), that is also provided by Christoffersen et al. (2013), could be conducted. A number of different stochastic volatility models (see Shephard & Andersen (2009) for an overview) could be used to test if the finding of this thesis of decreasing variance premium with maturity, is robust to model use.

In an incomplete market, option positions cannot always be perfectly hedged, resulting in a jump-based risk premium (see Hsieh et al. (2018)). Possibly an increase of the hedging premium on longer maturity could explain the strong increase in implied volatilities with maturity. Research into the role that hedging errors play in the increase of implied volatilities with the maturity horizon would give crucial information on the applicability of the results of this thesis.

Miron & Tudor (2010) find that GARCH models with normal errors are not able to fully comprehend the level of skewness as is present in the market, and that other error distributions are preferred. As a further generalisation of the innovation distribution, Barone-Adesi et al. (2008) propose option pricing method under asymmetric GARCH based on historical observations with more flexibility. Non-parametric methods are applied for the innovation distribution and the pricing kernel. This more flexible approach could prove to be an enhancement in a long maturity context as well.

[a] Possibly as a consequence of liquidity issues the implied volatilities are higher on longer maturity due to a hedging risk premium, see the further research suggestions in 6.1.

A. $T > 10$ scenario calibration and comparison to Ortec Finance Risk-Neutral model

The standard ($\xi = 0$) HNG-model is calibrated under $\mathbb{P}$ to sets of economic scenarios on a 20 and 30 years horizon. The results are compared to results from the Ortec Finance Risk-Neutral model (see Ortec Finance (2018)), a version of the continuous time Heston (1993) stochastic volatility model. This model was calibrated under $\mathbb{Q}$, against options of varying maturity, up to ten years. First in figure A.1 a plot is made of the Heston (1993) IV's, together with a plot of the HNG-model that is calibrated on 20/30 year scenarios without any premium (i.e. $\xi = 0$).

Next, in figure A.2 the same Heston (1993) model IV's are plotted with IV's from the HNG-model where the variance premium parameter ξ is such that the ξ-factor $(1 - 2\hat{\alpha}\hat{\xi})^{-1}$ as was estimated for $T = 10$ is kept constant for $T = 20y/T = 30y$. That is, for $T = 20y$, with $\hat{\alpha}^T$ the α as estimated for maturity T, and $\hat{\xi}^T$ similar

$$\xi^{20} = \{\xi \in \mathbb{R} : \frac{1}{1 - 2\hat{\alpha}^{20}\xi} = \frac{1}{1 - 2\hat{\alpha}^{10}\hat{\xi}^{10}}\} = \frac{\hat{\alpha}^{10}}{\hat{\alpha}^{20}}\xi^{10} \tag{A.1}$$

and in the same manner ξ^{30} is computed.

In figure A.3 a similar plot is made, where the variance premium was extrapolated. That is, the near linear increase in ξ-factor $(1-2\alpha\xi)^{-1}$ was extrapolated to 20/30 year horizon, and this was then solved to ξ^{20} and ξ^{30} similar to (A.1). It is seen that this extrapolation procedure results in extremely large implied volatilities, as could have been expected.

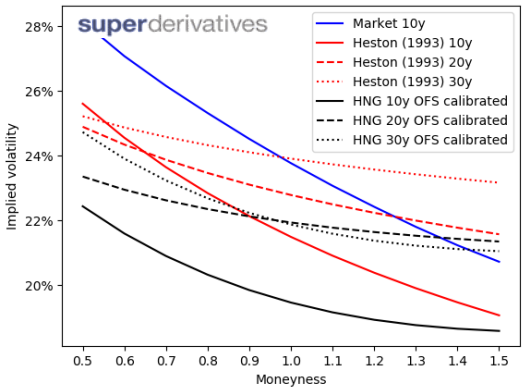

Figure A.1. IV-skew of HNG-model with 20/30 year scenario calibration with $\xi = 0$ and of the OF Risk Neutral Heston (1993)-model

Source market data: Superderivatives.

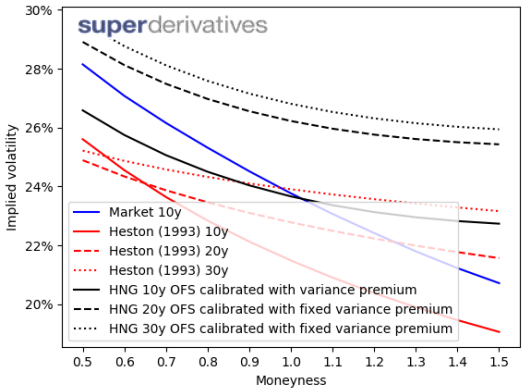

Figure A.2. IV-skew of HNG-model with 20/30 year scenario calibration with fixed variance premium, and of the OF Risk Neutral Heston (1993)-model

Source market data: Superderivatives.

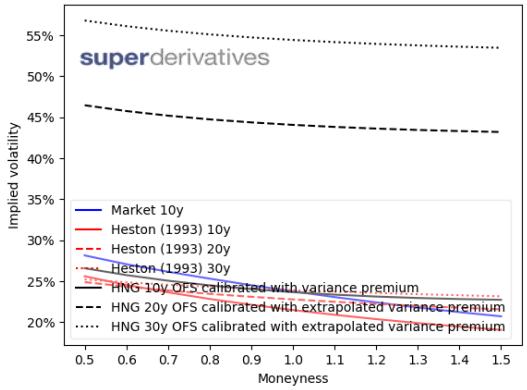

Figure A.3. IV-skew of HNG-model with 20/30 year scenario calibration with extrapolated variance premium, and of the OF Risk Neutral Heston (1993)-model

Source market data: Superderivatives.

B. Tables

Table B.1. Daily HNG-parameters for $T = 10$

	$\hat{\theta}_{hist}$	$\hat{\theta}^{*NTM}_{seq}$, $\hat{\theta}^{*AFTM}_{seq}$,	$\hat{\theta}^{*NTM}_{ad-hoc}$, $\hat{\theta}^{*AFTM}_{ad-hoc}$,		
$\hat{\omega}$	1.56e-11	1.56e-11	1.56e-11	·	·
$\hat{\alpha}$	4.01e-06	4.01e-06	4.01e-06	·	·
$\hat{\beta}$	0.819	0.819	0.819	·	·
$\hat{\gamma}$	189	189	189	·	·
$\hat{\lambda}$	2.23	2.23	2.23	·	·
h_{eq}	0.163	0.163	0.163	·	·
$\hat{\xi}$	·	3.42e+04	3.51e+04	·	·
$(1 - 2\hat{\alpha}\hat{\xi})^{-1}$	·	1.38	1.39	·	·
$\hat{\omega}^*$	·	2.16e-11	2.18e-11	-8.02e-08	-2.81e-07
$\hat{\alpha}^*$	·	7.61e-06	7.77e-06	3.68e-07	1.61e-06
$\hat{\beta}^*$	·	0.819	0.819	0.625	1.94e-08
$\hat{\gamma}^*$	·	139	138	1.01e+03	786
h^*_{eq}	·	0.239	0.242	0.271	0.271
LL	2.23e+04	-15.2	-253	0.412	-2.57
RMSE	·	0.804	2.18	0.641	0.615

An overview of the fitted daily HNG-parameters. The following constraints are imposed: $\hat{\omega} + \hat{\alpha} > 0$ is put in place to ensure positive initial volatility, and non-negative price of risk $\hat{\lambda} \geq -\frac{1}{2}$ was restricted (from the fact that the risk-neutralised version of the standard ($\mathbb{P}$-calibrated and with $\xi = 0$) model has $\hat{\lambda} = -\frac{1}{2}$, as can be seen from (4.6), it is found that non-negative price of risk restriction corresponds to setting $\hat{\lambda} \geq -\frac{1}{2}$). The first column of the table gives historically calibrated parameters $\hat{\theta}_{hist}$, with corresponding annualised equilibrium volatility h_{eq} (4.12) and with total log likelihood in the last row, as estimated against log returns of S&P 500 between 1990 and 2016, with historical interest rates. The second and third column show the sequential (based on $\hat{\theta}_{hist}$) $\mathbb{Q}$-parameter estimations for Near The Money (NTM, with relative moneyness m between 0.9 and 1.1, that is strike price $0.9 \cdot S_t \leq K \leq 1.1 \cdot S_t$ with S_t current asset value of the underlying) and Away From The Money (AFTM, $0.5 \leq m \leq 1.5$) options with maturity T equal to 10 years. The annualised risk neutral equilibrium volatility h^*_{eq} is also displayed, as well as the root mean squared error (RMSE) of model IV's with respect to market IV's.

Table B.2. Monthly HNG-parameters for $T = 10$

	$\hat{\theta}^{M}_{hist}$	$\hat{\theta}^{M}_{scen}$	$\hat{\theta}^{*M}_{seq}$	$\hat{\theta}^{*M}_{scen}$	$\hat{\theta}^{*M}_{ad-hoc}$
$\hat{\omega}$	1.67e-05	0.00075	1.67e-05	0.00075	·
$\hat{\alpha}$	0.000701	0.000239	0.000701	0.000239	·
$\hat{\beta}$	0.485	0.413	0.485	0.413	·
$\hat{\gamma}$	0.000203	16.8	0.000203	16.8	·
$\hat{\lambda}$	3.1	1.42	3.1	1.42	·
h_{eq}	0.129	0.151	0.129	0.151	·
$\hat{\xi}$	·	·	329	1.04e+03	·
$(1 - 2\hat{\alpha}\hat{\xi})^{-1}$	·	·	1.86	1.98	·
$\hat{\omega}^*$	·	·	3.1e-05	0.00149	-3.36e-05
$\hat{\alpha}^*$	·	·	0.00242	0.000937	5.55e-06
$\hat{\beta}^*$	·	·	0.485	0.413	0.397
$\hat{\gamma}^*$	·	·	2.17	9.69	331
h^*_{eq}	·	·	0.242	0.241	0.25
LL	1.39e+04	4.66e+05	-277	-269	-95.4
RMSE	·	·	2.45	2.36	0.985

Table B.3. Yearly HNG-parameters for $T = 10$

	$\hat{\theta}^{Y}_{hist}$	$\hat{\theta}^{Y}_{scen}$	$\hat{\theta}^{*Y}_{seq}$	$\hat{\theta}^{*Y}_{scen}$	$\hat{\theta}^{*Y}_{ad-hoc}$
$\hat{\omega}$	0	0.00274	0	0.00274	·
$\hat{\alpha}$	0.00872	0.0304	0.00872	0.0304	·
$\hat{\beta}$	0.539	0.0842	0.539	0.0842	·
$\hat{\gamma}$	3.27e-09	2.28	3.27e-09	2.28	·
$\hat{\lambda}$	2.48	-0.428	2.48	-0.428	·
h_{eq}	0.137	0.209	0.137	0.209	·
$\hat{\xi}$	·	·	22.9	3.05	·
$(1 - 2\hat{\alpha}\hat{\xi})^{-1}$	·	·	1.66	1.23	·
$\hat{\omega}^*$	·	·	0	0.00336	-0.00773
$\hat{\alpha}^*$	·	·	0.0241	0.0457	0.0119
$\hat{\beta}^*$	·	·	0.539	0.0842	0.691
$\hat{\gamma}^*$	·	·	1.99	2.01	4.59
h^*_{eq}	·	·	0.257	0.259	0.272
LL	5.05e+03	2.8e+04	-186	-186	-1.79
RMSE	·	·	1.55	1.47	0.612

The monthly and yearly formulations of the HNG-model are estimated under $\mathbb{P}$: This is done against historical S&P 500 log returns between 1990 and 2016 ($\hat{\theta}^{f}_{hist}$ with frequency $f \in \{M, Y\}$) with historical interest rates, and against 10 year scenarios of US equity total returns ($\hat{\theta}^{f}_{scen}$), where corresponding 3 month swap rate scenarios are used. A rolling window approach is applied. Parameters are sequentially calibrated under $\mathbb{Q}$ against options with maturity of ten years with relative moneyness ranging between 0.5 and 1.5. The same constraints as before ($\hat{\omega} + \hat{\alpha} > 0$, $\hat{\lambda} \geq -\frac{1}{2}$) are imposed.

C. Figures

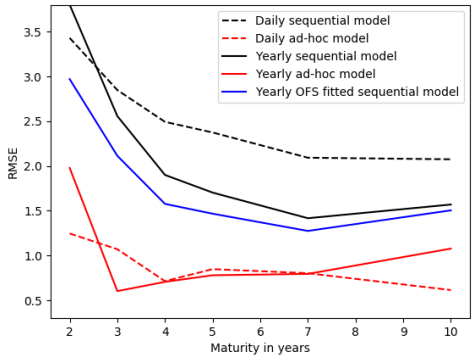

Figure C.1. Out of sample away from the money RMSE

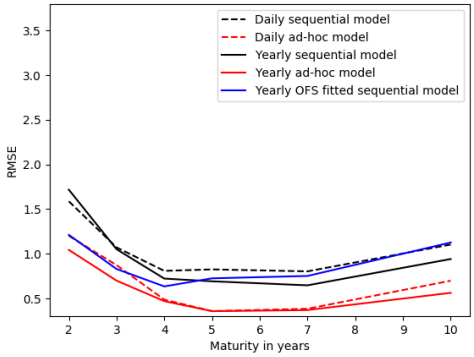

Figure C.2. Near the money RMSE

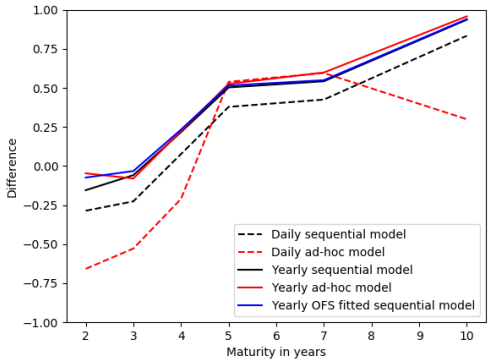

Figure C.3. Away from the money bias

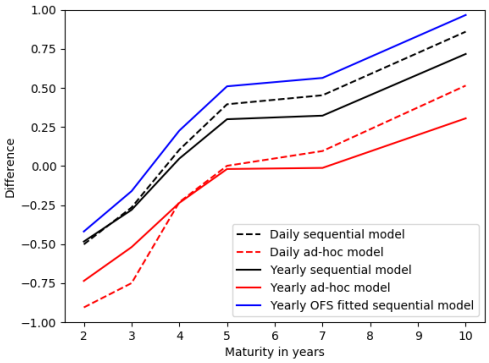

Figure C.4. Near the money bias

The top two figures show the in sample RMSE of the model IV with respect to the market IV, plotted against maturity in years T for the sequential and ad-hoc HNG-models with varying frequency. The bottom two figures show the mean in sample bias, that is $IV^{Market} - IV^{HNG}$. The out of sample period ranges from 2019-01 to 2020-02.

D. Technicalities

The asset dynamics assumption from section 4.1 is restated:

Assumption. The total return asset time series $(S_t)_{t\in[0,T]}$ evolves discretely with time steps Δ by the following process:

$$\begin{aligned} \ln\left(\frac{S_t}{S_{t-\Delta}}\right) &= r_{t-\Delta} + \lambda h_t + \sqrt{h_t} z_t \\ h_t &= \omega + \beta h_{t-\Delta} + \alpha\big(z_{t-\Delta} - \gamma\sqrt{h_{t-\Delta}}\big)^2, \end{aligned} \tag{D.1}$$

with r_t the continuously compounded interest rate per time step, that is known at t and constant on time interval $[t, t+\Delta]$, and where z_t is a standard normal innovation term conditional on $\mathcal{F}_{t-\Delta}$. The quantity h_t is the conditional volatility of the log return between $t-\Delta$ and t, which is predictable in the sense that it is $\mathcal{F}_{t-\Delta}$-measurable.

D.1. Covariance

The $\mathcal{F}_{t-\Delta}$-conditional covariance between spot return and volatility is computed:

$$\begin{aligned} \mathrm{Cov}_{t-\Delta}(h_{t+\Delta}, \ln(S_t)) &= \mathbb{E}_{t-\Delta}\left[(h_{t+\Delta} - \mathbb{E}_{t-\Delta}[h_{t+\Delta}])(\ln(S_t) - \mathbb{E}_{t-\Delta}[\ln(S_t)])\right] \\ &= \mathbb{E}_{t-\Delta}\left[(\alpha z_t^2 - 2\alpha\gamma\sqrt{h_t} z_t + \alpha\gamma^2 h_t - \alpha)(\sqrt{h_t} z_t)\right] \\ &= \mathbb{E}_{t-\Delta}[\alpha\sqrt{h_t} z_t^3 - 2\alpha\gamma h_t z_t^2 + \alpha\gamma^2 h_t \sqrt{h_t} z_t - \alpha\sqrt{h_t} z_t] \\ &= -2\alpha\gamma h_t. \end{aligned}$$

D.2. Dividend correction

The dividend correction (section 4.3.2) is intuitively justified as follows: Consider a vanilla call C on some asset that is currently at S_t. A replicating portfolio contains $\frac{\partial C}{\partial S_t}$ units of this asset. This replicating portfolio must have the same price as C by arbitrage arguments (see for example chapter 5 of Damodaran (2002) for theory on replicating portfolios). If this asset were to pay out a relative amount q in dividends (present value) over the maturity of the option, this would lead to an increase of value of the replicating portfolio of $qS_t\frac{\partial C}{\partial S_t}$. If on the other hand this asset were to not pay out dividends, but instead it would increase in value by qS_t, from S_t to $(1+q)S_t$, this would lead to an increase of the call value of approximately (for small q) $qS_t\frac{\partial C}{\partial S_t}$, so to the same increase in the value of the option. We end up with the same

value for C for the non-dividend paying asset that is currently at $S_t(1+q)$ as the underlying, as with the dividend paying asset that is at S_t, i.e. $S_t(1+q)$ minus the current value of the dividend pay-out during lifetime of the option as underlying.

D.3. Heston (1993) as continuous limit

In Heston & Nandi (2000) it is demonstrated that (D.1) converges to a version of the Heston (1993) stochastic volatility model, if it is assumed that the volatility per time unit has a well-defined continuous-time limit. A more detailed proof is outlined here. The diffusion limit will follow as a consequence of the following theorem, as stated by Nelson et al. (1994):

Theorem 1 (Stroock-Varadhan)**.** *A stochastic integral equation on the filtered probability space* $(\Omega, \mathcal{F}_T, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P})$ *of the following form is given:*

$$X_t = X_0 + \int_0^t \mu(X_s)\,ds + \int_0^t \sigma(X_s)\,dW_s \tag{D.2}$$

where W_s *is a standard Brownian motion and with* $\mu(x)$ *and* $\sigma(x)$ *continuous functions on* $\mathbb{R}$ *where* X_0 *has initial cumulative distribution* $F(x)$*. Assume that* (D.2) *has a weak-sense solution*[a]*. Suppose there exists some discrete time Markov process* $(Y_t)_{t\in\mathbb{N}}$ *with initial value* Y_0 *with cumulative distribution* $\mathbb{P}[Y \le x] = F_Y(x)$*, and that for*

$$\mu_\Delta = \frac{\mathbb{E}[Y_{t+1} - Y_t | Y_t = y]}{\Delta} \tag{D.3}$$

$$\sigma_\Delta = \frac{\mathrm{Var}[Y_{t+1} - Y_t | Y_t = y]}{\Delta}, \tag{D.4}$$

the following conditions are satisfied for some $\delta > 0$*:*

(i) $F_Y(x) \overset{\Delta\to 0}{\to} F(x)$,

(ii) $\mu_\Delta(x) \overset{\Delta\to 0}{\to} \mu(x)$,

(iii) $\sigma_\Delta(x) \overset{\Delta\to 0}{\to} \sigma(x)$,

(iv) $\frac{\mathbb{E}[|Y_{t+1}-Y_t|^{2+\delta}|Y_t=y]}{\Delta} \overset{\Delta\to 0}{\to} 0$

where (i) holds for all continuity points of F*, and (ii)-(iv) hold uniformly. Then on any finite interval* $[0,T]$*, the process* $(X_t^\Delta)_{t\in[0,T]} := Y_{\lfloor t/\Delta \rfloor}$ *(where* $\lfloor x \rfloor$ *denotes the nearest integer below* x*) converges weakly to a solution of* (D.2)*.*

Proof. See Appendix of Nelson et al. (1994).

[a] See appendix A of Nelson (1990) for sufficient conditions for the existence of a weak-sense solution.

In the Heston (1993) process, the volatility follows a CIR (Cox et al. 1985) process:

$$v_t = v_0 + \int_0^t \kappa(\theta - v)\mathrm{d}t + \int_0^t \sigma\sqrt{v}\mathrm{d}W. \tag{D.5}$$

It is claimed that all conditions of Stroock-Varadhan are satisfied for discrete time Markov process $(v_t^\Delta)_{t\in[0,\Delta,2\Delta,\ldots,T]} := h_{t\in[0,\Delta,2\Delta,\ldots,T]}/\Delta$, for some deterministic initial value $v_0 \in \mathbb{R}$. As was mentioned at the beginning of this appendix, it is assumed that v_t^Δ has a well-defined continuous-time limit. Taking v_0 deterministic immediately gives condition (i). It is also clear that $(v_t^\Delta)_{t\in[0,\ldots,T]}$ is a discrete time Markov-process, as v_t^Δ is recursively defined from v_{t-1}^Δ.

Equation (D.1) gives:

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

so that (using predictability of h_t):

$$\begin{aligned} \mathbb{E}_{t-\Delta}[h_{t+\Delta}] &= \omega + \beta h_t + \alpha\mathbb{E}_{t-\Delta}[z_t^2] + \alpha\gamma^2 h_t - 2\alpha\gamma\sqrt{h_t}\mathbb{E}_{t-\Delta}[z_t] \\ &= \omega + \alpha + (\beta + \alpha\gamma^2)h_t \end{aligned}$$

and

$$\begin{aligned} \mathrm{Var}_{t-\Delta}[h_{t+\Delta}] &= \mathrm{Var}_{t-\Delta}[\alpha z_t^2 - 2\alpha\gamma\sqrt{h_t}z_t] \\ &= \alpha^2(2 + 4\gamma^2 h_t) \end{aligned} \tag{D.7}$$

where it was used that z_t^2 has a χ_1^2 distribution with variance 2, and that

$$\mathrm{Cov}[z_t^2, z_t] = \mathbb{E}[z_t^3] - \mathbb{E}[z_t]\mathbb{E}[z_t^2] = 0.$$

For the variance per time unit ($v_t^\Delta = h_t/\Delta$) (D.6) becomes:

$$v_{t+\Delta}^\Delta = \frac{\omega}{\Delta} + \beta v_t^\Delta + \frac{\alpha}{\Delta}(z_t - \gamma\sqrt{\Delta}\sqrt{v_t^\Delta})^2, \tag{D.8}$$

so that

$$\mathbb{E}_{t-\Delta}[v_{t+\Delta}^\Delta] = \frac{\omega}{\Delta} + \beta v_t^\Delta + \frac{\alpha}{\Delta} + \alpha\gamma^2 v_t^\Delta.$$

Equation (D.8) is transformed by setting

$$\begin{aligned} \omega(\Delta) &= (\kappa\theta - \frac{1}{4}\sigma^2)\Delta^2 \\ \alpha(\Delta) &= \frac{1}{4}\sigma^2\Delta^2 \\ \beta(\Delta) &= 0 \\ \gamma(\Delta) &= \frac{2}{\sigma\Delta} - \frac{\kappa}{\sigma} \\ \lambda(\Delta) &= \lambda \end{aligned} \tag{D.9}$$

so that it becomes

$$v_{t+\Delta}^{\Delta} = \Delta(\kappa\theta - \frac{1}{4}\sigma^2) + \frac{1}{4}\sigma^2\Delta\left(z_t - \left(\frac{2}{\sigma\sqrt{\Delta}} - \frac{\sqrt{\Delta}\kappa}{\sigma}\right)\sqrt{v_t^{\Delta}}\right)^2. \tag{D.10}$$

It is obtained that

$$\begin{aligned}\mathbb{E}_{t-\Delta}[v_{t+\Delta}^{\Delta} - v_t^{\Delta}] &= (\kappa\theta - \frac{1}{4}\sigma^2)\Delta + \frac{1}{4}\sigma^2\Delta + \\ &\qquad \left(\frac{1}{4}\sigma^2\Delta^2(2/(\sigma\Delta) - \kappa/\sigma)^2\right)v_t^{\Delta} - v_t^{\Delta} \\ &= \left(\kappa\theta - \kappa v_t^{\Delta}\right)\Delta + \frac{1}{4}\kappa^2\Delta^2 v_t^{\Delta} \end{aligned} \tag{D.11}$$

and from (D.7) it follows that

$$\begin{aligned}\mathrm{Var}_{t-\Delta}[v_{t+\Delta}^{\Delta}] &= \frac{\mathrm{Var}_{t-\Delta}[h_{t+\Delta}]}{\Delta^2} \\ &= \frac{\frac{1}{16}\sigma^4\Delta^4(2 + 4(2/(\sigma\Delta) - \kappa/\sigma)^2)v_t^{\Delta}\Delta}{\Delta^2} \\ &= \sigma^2\Delta v_t^{\Delta} + \left(\frac{1}{8}\sigma^4 v_t^{\Delta}\Delta + \frac{\kappa^2\Delta v_t^{\Delta}}{4\sigma^2} - \frac{\kappa v_t^{\Delta}}{\sigma^2}\right)\Delta^2 \end{aligned} \tag{D.12}$$

and by predictability of v_t^{Δ}

$$\mathrm{Var}_{t-\Delta}[v_{t+\Delta}^{\Delta} - v_t^{\Delta}] = \mathrm{Var}_{t-\Delta}[v_{t+\Delta}^{\Delta}]. \tag{D.13}$$

By predictability of v_t^{Δ}, it follows from (D.11) that

$$\frac{\mathbb{E}[v_{t+1}^{\Delta} - v_t^{\Delta} | v_t^{\Delta} = v]}{\Delta} = \frac{\mathbb{E}_{t-\Delta}[v_{t+\Delta}^{\Delta} - v_t^{\Delta}]}{\Delta} \overset{\Delta\to 0}{\to} \kappa\theta - \kappa v_t^{\Delta}$$

and (D.12) and (D.13) give that

$$\frac{\mathrm{Var}[v_{t+1}^{\Delta} - v_t^{\Delta} | v_t^{\Delta} = v]}{\Delta} = \frac{\mathrm{Var}_{t-\Delta}[v_{t+\Delta}^{\Delta} - v_t^{\Delta}]}{\Delta} \overset{\Delta\to 0}{\to} \sigma^2 v_t^{\Delta}$$

so that conditions (ii) and (iii) of Stroock-Varadhan are satisfied. The fourth (absolute) moment of $v_{t+\Delta}^{\Delta} - v_t^{\Delta}$ is computed in order to confirm that condition (iv) holds for $\delta = 2$:

$$\begin{aligned}\mathbb{E}_{t-\Delta}[|v_{t+1}^{\Delta} - v_t^{\Delta}|^4] &= \mathbb{E}_{t-\Delta}[(v_{t+1}^{\Delta} - v_t^{\Delta})^4] \\ &= \mathbb{E}_{t-\Delta}\Bigg[\Bigg(\Delta(\kappa\theta - \frac{1}{4}\sigma^2) + \\ &\qquad \frac{1}{4}\sigma^2\Delta\Big(z_t - \Big(\frac{2}{\sigma\sqrt{\Delta}} - \frac{\sqrt{\Delta}\kappa}{\sigma}\Big)\sqrt{v_t^{\Delta}}\Big)^2\Bigg)^4\Bigg]. \end{aligned}$$

By tedious algebraic reordering this can be rewritten to:

$$\begin{aligned}\frac{1}{256}\Delta^2\mathbb{E}_{t-\Delta}\Big[\Big(\Delta\sqrt{\Delta}\kappa^2 v_t^{\Delta} - 2\sqrt{\Delta}\kappa\Big(\sqrt{v_t^{\Delta}}\Big(2\sqrt{v_t^{\Delta}} \\ - \sqrt{\Delta}\sigma z_t\Big) - 2\theta\Big) - \sigma\Big(4\sqrt{v_t^{\Delta}}z_t - \sqrt{\Delta}\sigma(z_t^2 - 1)\Big)\Big)^4\Big]. \end{aligned} \tag{D.14}$$

As all terms in (D.14) are $\mathcal{O}(\Delta^n)$ with $n \geq 2$, by the assumption that $(v_t^\Delta)_{t\in[0,T]}$ has a well-defined limit, the vact that the standard normal z_t has bounded first to fourth moments, and as all other terms are constants, it is confirmed that

$$\frac{\mathbb{E}_{t-\Delta}\left[|v_{t+\Delta}^\Delta - v_t^\Delta|^4\right]}{\Delta} \overset{\Delta\to 0}{\to} 0.$$

With conditions (i)-(iv) satisfied, Stroock-Varadhan gives that the process $\tilde{v}_t^\Delta := (v_{[t/\Delta\rfloor}^\Delta)_{t\in[0,T]}$ converges weakly to the solution of (D.5), or equivalently, the dynamics of the diffusion limit are:

$$\mathrm{d}v = \kappa(\theta - v)\mathrm{d}t + \sigma\sqrt{v}\mathrm{d}W$$

with W a Brownian motion. Note that this diffusion limit is the limit of the process (D.10), with skewness parameter $\sqrt{\Delta}\gamma(\Delta)$. By similar, more straightforward arguments, it is seen that the Stroock-Varadhan theorem can be applied to the log return process $(\ln(S_t))_{t\in[0,T]}$ as well, so that the dynamics of the diffusion limit are:

$$\mathrm{d}\ln S = (r + \lambda v)\mathrm{d}t + \sqrt{v}\mathrm{d}\hat{W}$$

for a Brownian motion $\hat{W}$. It is claimed that $W = -\hat{W}$. This follows if $\mathcal{F}_{t-\Delta}$-conditional correlation between the variance process and the continuously compounded stock return is equal to -1, i.e.

$$\mathrm{Corr}[v_{t+\Delta}^\Delta, \ln S_t] = \frac{\mathrm{Cov}[v_{t+\Delta}^\Delta, \ln S_t]}{\sqrt{\mathrm{Var}_{t-\Delta}[v_{t+\Delta}^\Delta]}\sqrt{\mathrm{Var}_{t-\Delta}[\ln S_t]}} \overset{\Delta\to 0}{\to} -1.$$

To show this,

$$v_{t+\Delta}^\Delta - \mathbb{E}_{t-\Delta}[v_{t+\Delta}^\Delta] = \frac{\alpha(\Delta)}{\Delta}\left(z_t^2 - 2z_t\gamma(\Delta)\sqrt{\Delta}\sqrt{v_t^\Delta} - 1\right)$$

is used to find

$$\begin{aligned}
\mathrm{Cov}_{t-\Delta}[v_{t+\Delta}^\Delta, \ln(S_t)] &= \mathbb{E}_{t-\Delta}\left[(v_{t+\Delta}^\Delta - \mathbb{E}_{t-\Delta}[v_{t+\Delta}^\Delta])(\ln(S_t) - \mathbb{E}_{t-\Delta}[\ln(S_t)])\right] \\
&= \mathbb{E}_{t-\Delta}\left[\frac{\alpha(\Delta)}{\Delta}\left(z_t^2 - 2z_t\gamma(\Delta)\sqrt{\Delta}\sqrt{v_t^\Delta} - 1\right)\sqrt{v_t^\Delta}\sqrt{\Delta}z_t\right] \\
&= -2\alpha(\Delta)\gamma v_t^\Delta
\end{aligned}$$

so that (writing $(\alpha(\Delta))^2$ as $\alpha^2(\Delta)$ and similar for γ):

$$\begin{aligned}
\mathrm{Corr}_{t-\Delta}[v_{t+\Delta}^\Delta, \ln(S_t)] &= \frac{-2\alpha(\Delta)\gamma(\Delta)v_t^\Delta}{\sqrt{\mathrm{Var}_{t-\Delta}[\ln(S_t)]}\sqrt{\mathrm{Var}_{t-\Delta}[v_{t+\Delta}^\Delta]}} \\
&= \frac{-2\sqrt{\Delta}\alpha(\Delta)\gamma(\Delta)v_t^\Delta}{\sqrt{v_t^\Delta}\sqrt{\alpha^2(\Delta)(2 + 4\gamma(\Delta))^2 v_t^\Delta\Delta)}} \\
&= \frac{-\sqrt{2\gamma^2(\Delta)\Delta v_t^\Delta}}{\sqrt{1 + 2\gamma^2(\Delta)\Delta v_t^\Delta}}.
\end{aligned} \tag{D.15}$$

By $\gamma^2(\Delta)\Delta \overset{\Delta\to 0}{\to} \infty$ it follows that (D.15) goes to -1. This shows that $W = -\hat{W}$, so that the HNG-process converges to the CIR (Cox et al. 1985), Heston (1993) stochastic volatility model:

$$\begin{aligned} \mathrm{d}\ln(S) &= (r + \lambda v)\mathrm{d}t + \sqrt{v}\mathrm{d}W \\ \mathrm{d}v &= \kappa(\theta - v)\mathrm{d}t - \sigma\sqrt{v}\mathrm{d}W \end{aligned}$$

with W a Brownian motion.

D.4. Moment generating function

This section contains details of the derivation of the recursive formula for the moment generating function (4.9) by Heston & Nandi (2000).

Define

$$f(t; T, \phi) = \mathbb{E}_t[S_T^\phi] = \mathbb{E}_t[\exp(\phi X_T)] \tag{D.16}$$

with

$$X_t = \ln(S_t).$$

Note that $f(i\phi)$ (or $\tilde{f}(\phi) := f(i\phi)$, to be precise) is the t-conditional characteristic function of X_T. A log-linear form of $f(t; T, \phi)$ is stipulated:

$$f(t; T, \phi) = \exp\left(\phi X_t + A(t; T, \phi) + B(t; T, \phi)h_{t+\Delta}\right). \tag{D.17}$$

By $\mathcal{F}_T$-measurability, it holds that

$$f(T; T, \phi) = \exp(\phi X_T),$$

which gives the following terminal condition:

$$A(T; T, \phi) = B(T; T, \phi) = 0.$$

By the tower property

$$\begin{aligned} f(t; T, \phi) &= \mathbb{E}_t[f(t + \Delta; T, \phi)] \\ &= \mathbb{E}_t[\exp(\phi X_{t+\Delta} A(t + \Delta; T, \phi) + B(t + \Delta; T, \phi)h_{t+2\Delta})] \end{aligned}$$

and plugging in GARCH dynamics (D.1) gives

$$\begin{aligned} f(t; T, \phi) = \mathbb{E}_t\Bigg[\exp\Bigg(\phi\Big(X_t + r + \lambda h_{t+\Delta} + \sqrt{h_{t+\Delta}} z_{t+\Delta} \Big) + \\ A(t + \Delta; T, \phi) + B(t + \Delta; T, \phi) \times \\ \Big(\beta h_{t+\Delta} + \alpha\big(z_{t+\Delta} - \gamma\sqrt{h_{t+\Delta}} \big)^2 + \omega \Big) \Bigg) \Bigg]. \end{aligned} \tag{D.18}$$

This expression can be rearranged to (denoting $A(t;T,\phi)$ as A_t, and $B(t;T,\phi)$ as B_t):

$$f(t;T,\phi) = \mathbb{E}_t\Big[\exp\Big(\phi(X_t+r) + A_{t+\Delta} + B_{t+\Delta}\omega + B_{t+\Delta}\alpha\left(z_{t+\Delta} - \Big(\gamma - \frac{\phi}{2B_{t+\Delta}\alpha}\Big)\sqrt{h_{t+\Delta}}\right)^2 + \Big(\phi\lambda + B_{t+\Delta}\beta + \phi\gamma - \frac{\phi^2}{4B_{t+\Delta}\alpha}\Big)h_{t+\Delta}\Big)\Big]. \tag{D.19}$$

The exponential term in (D.19) is rewritten:

$$\begin{aligned}
&\phi(X_t+r) + A_{t+\Delta} + B_{t+\Delta}\omega + B_{t+\Delta}\alpha\left(z_{t+\Delta} - \Big(\gamma - \frac{\phi}{2B_{t+\Delta}\alpha}\Big)\sqrt{h_{t+\Delta}}\right)^2 + \\
&\qquad \Big(\phi\lambda + B_{t+\Delta}\beta + \phi\gamma - \frac{\phi^2}{4B_{t+\Delta}\alpha}\Big)h_{t+\Delta} \\
&= B_{t+\Delta}\alpha\Big(z_{t+\Delta} - \left(\gamma - \frac{\phi}{2B_{t+\Delta}\alpha}\right)\sqrt{h_{t+\Delta}}\Big)^2 - \frac{\phi^2}{4B_{t+\Delta}\alpha}h_{t+\Delta} + \\
&\qquad \phi\gamma h_{t+\Delta} + \phi(X_t+r) + A_{t+\Delta} + B_{t+\Delta}\omega + h_{t+\Delta}(\phi\lambda + B_{t+\Delta}\beta) \\
&= B_{t+\Delta}\alpha\left(z_{t+\Delta}^2 + h_{t+\Delta}\left(\gamma - \frac{\phi}{2B_{t+\Delta}\alpha}\right)^2 - 2z_{t+\Delta}\left(\gamma - \frac{\phi}{2B_{t+\Delta}\alpha}\right)\sqrt{h_{t+\Delta}}\right) - \\
&\qquad \frac{\phi^2}{4B_{t+\Delta}\alpha}h_{t+\Delta} + \phi\gamma h_{t+\Delta} + \phi(X_t+r) + A_{t+\Delta} + B_{t+\Delta}\omega + h_{t+\Delta}(\phi\lambda + B_{t+\Delta}\beta) \\
&= B_{t+\Delta}\alpha z_{t+\Delta}^2 + B_{t+\Delta}\alpha h_{t+\Delta}\gamma^2 + h_{t+\Delta}\frac{\phi^2}{4B_{t+\Delta}\alpha} - \gamma h_{t+\Delta}\phi - \\
&\qquad 2B_{t+\Delta}\alpha z_{t+\Delta}\gamma\sqrt{h_{t+\Delta}} + z_{t+\Delta}\phi\sqrt{h_{t+\Delta}} - \frac{\phi^2}{4B_{t+\Delta}\alpha}h_{t+\Delta} + \phi\gamma h_{t+\Delta} + \\
&\qquad \phi(X_t+r) + A_{t+\Delta} + B_{t+\Delta}\omega + h_{t+\Delta}(\phi\lambda + B_{t+\Delta}\beta) \\
&= B_{t+\Delta}\alpha z_{t+\Delta}^2 + B_{t+\Delta}\alpha h_{t+\Delta}\gamma^2 - 2B_{t+\Delta}\alpha z_{t+\Delta}\gamma\sqrt{h_{t+\Delta}} + z_{t+\Delta}\phi\sqrt{h_{t+\Delta}} + \\
&\qquad \phi(X_t+r) + A_{t+\Delta} + B_{t+\Delta}\omega + h_{t+\Delta}(\phi\lambda + B_{t+\Delta}\beta) \\
&= \phi\sqrt{h_{t+\Delta}}z_{t+\Delta} + \alpha B_{t+\Delta}(z_{t+\Delta} - \gamma\sqrt{h_{t+\Delta}})^2 + \\
&\qquad \phi(X_t+r) + A_{t+\Delta} + B_{t+\Delta}\omega + h_{t+\Delta}(\phi\lambda + B_{t+\Delta}\beta).
\end{aligned}$$

Using predictability of h_t, (D.19) can be written as

$$f(t;T,\phi) = \mathbb{E}_t\left[\exp\left(\alpha B_{t+\Delta}\left(z_{t+\Delta} - \Big(\gamma - \frac{\phi}{2B_{t+\Delta}\alpha}\Big)\sqrt{h_{t+\Delta}}\right)^2\right)\right] \times \exp\left(\phi(X_t+r) + A_{t+\Delta} + B_{t+\Delta}\omega + \Big(\phi(\lambda+\gamma) + B_{t+\Delta}\beta - \frac{\phi^2}{4B_{t+\Delta}\alpha}\Big)h_{t+\Delta}\right). \tag{D.20}$$

The following result for $z \sim \mathcal{N}(0,1)$ will be used:

$$\mathbb{E}[\exp(a(z+b)^2)] = \exp\left(-\frac{1}{2}\ln(1-2a) + \frac{ab^2}{1-2a}\right). \tag{D.21}$$

This holds only for $0 < a < \frac{1}{2}$, so to use this for (D.20) it is needed that $\alpha B_{t+\Delta} < \frac{1}{2}$. For all estimated parameters this is comfortably satisfied. Applying (D.21) to (D.20) gives:

$$\begin{aligned}
&\mathbb{E}_t\left[\exp\left(\alpha B_{t+\Delta}\left(z_{t+\Delta}-\left(\gamma-\frac{\phi}{2B_{t+\Delta}\alpha}\right)\sqrt{h_{t+\Delta}}\right)^2\right)\right] \\
&\quad=\exp\left(-\frac{1}{2}\ln(1-2\alpha B_{t+\Delta})+\frac{\alpha B_{t+\Delta}\left(\gamma-\frac{\phi}{2B_{t+\Delta}\alpha}\right)^2 h_{t+\Delta}}{1-2\alpha B_{t+\Delta}}\right) \\
&\quad=\exp\left(-\frac{1}{2}\ln(1-2\alpha B_{t+\Delta})+\frac{(\alpha B_{t+\Delta}\gamma^2+\frac{\phi^2}{4B_{t+\Delta}\alpha}-\gamma\phi)h_{t+\Delta}}{1-2\alpha B_{t+\Delta}}\right) \\
&\quad=\exp\left(-\frac{1}{2}\ln(1-2\alpha B_{t+\Delta})+\frac{(-\frac{1}{2}\gamma^2(1-2\alpha B_{t+\Delta})+\frac{1}{2}\gamma^2+\frac{\phi^2}{4B_{t+\Delta}\alpha}-\gamma\phi)h_{t+\Delta}}{1-2\alpha B_{t+\Delta}}\right) \\
&\quad=\exp\left(-\frac{1}{2}\gamma^2-\frac{1}{2}\ln(1-2\alpha B_{t+\Delta})+\frac{(\frac{1}{2}\gamma^2 h_{t+\Delta}+\frac{\phi^2}{4B_{t+\Delta}\alpha}-\gamma\phi)h_{t+\Delta}}{1-2\alpha B_{t+\Delta}}\right).
\end{aligned}$$

This is subsequently substituted into (D.19), which is equated to the stipulated log-linear form (D.17):

$$\begin{aligned}
&f(t;T,\phi)=\exp(\phi X_t+A_t+B_t h_{t+\Delta}) \qquad\qquad (D.22)\\
&\quad=\exp\left(-\frac{1}{2}\gamma^2 h_{t+\Delta}-\frac{1}{2}\ln(1-2\alpha B_{t+\Delta})+\frac{(\frac{1}{2}\gamma^2+\frac{\phi^2}{4B_{t+\Delta}\alpha}-\gamma\phi)h_{t+\Delta}}{1-2\alpha B_{t+\Delta}}\right)\times \\
&\qquad\exp\left(\phi(X_t+r)+A_{t+\Delta}+B_{t+\Delta}\omega+\left(\phi(\lambda+\gamma)+\beta B_{t+\Delta}-\frac{\phi^2}{4B_{t+\Delta}\alpha}\right)h_{t+\Delta}\right) \\
&\quad=\exp\left(\left(\phi(\lambda+\gamma)-\frac{1}{2}\gamma^2+\frac{\frac{1}{2}(\phi-\gamma)^2}{1-2\alpha B_{t+\Delta}}+\beta B_{t+\Delta}\right)h_{t+\Delta}-\frac{1}{2}\ln(1-2\alpha B_{t+\Delta})+\right. \\
&\qquad\left.\phi(X_t+r)+A_{t+\Delta}+B_{t+\Delta}\omega\right)
\end{aligned}$$

where in the last step it was used that

$$\frac{\frac{\phi^2}{4B_{t+\Delta}\alpha}}{1-2\alpha B_{t+\Delta}}-\frac{\phi^2}{4\alpha B_{t+\Delta}}=\frac{\frac{1}{2}\phi^2}{1-2\alpha B_{t+\Delta}}.$$

Equating left and right side of (D.22) the following recursive solution is established:

$$\begin{aligned}
A(t;T,\phi)&=A_{t+\Delta}+\phi r+\omega B_{t+\Delta}-\frac{1}{2}\ln(1-2\alpha B_{t+\Delta}) \\
B(t;T,\phi)&=\phi(\lambda+\gamma)-\frac{1}{2}\gamma^2+\beta B_{t+\Delta}+\frac{1/2(\phi-\gamma)^2}{1-2\alpha B_{t+\Delta}},
\end{aligned}$$

so that the conditional moment generating function (D.16) can be recursively derived from terminal conditions $A(T,T;\phi)=B(T,T;\phi)=0$, as required.

D.5. Option pricing formula

This section contains details of the derivation of the closed-form option pricing formula (4.10) by Heston & Nandi (2000).

Section D.4 gives a recursive expression for the $\mathcal{F}_t$-conditional moment generating function

$$f(\phi) = \mathbb{E}_t[\exp(\phi X_T)] \tag{D.23}$$

with $X_T = \ln(S_T)$ the logarithm of strike asset price at maturity, where dependency on t, T for f is dropped from the notation. Denote the corresponding $\mathcal{F}_t$-conditional probability density function (pdf) of X_T by $g(x)$. Define

$$\hat{g}(x) := \exp(x)\frac{g(x)}{f(1)}.$$

As by (D.23) $f(1) = \mathbb{E}_t[\exp(X_T)]$ and $g(x)$ is a pdf, $\hat{g}(x)$ is a pdf as well. The moment generating function corresponding to this pdf is

$$\int_{-\infty}^{\infty} \exp(\phi x)\hat{g}(x)\mathrm{d}x = \frac{1}{f(1)}\int_{-\infty}^{\infty} \exp((\phi+1)x)g(x)\mathrm{d}x \tag{D.24}$$

$$= \frac{E_t[\exp((\phi+1)X_T)]}{f(1)} = \frac{f(\phi+1)}{f(1)} \tag{D.25}$$

so that the characteristic functions that correspond to $g(x)$ and $\hat{g}(x)$ are given by $f(i\phi)$ and $\frac{f(i\phi+1)}{f(1)}$ respectively. As $S_T = \exp(X_T)$ the expectation of the pay-off for a call option with strike K at t is equal to (writing $\max(x, 0)$ as $(x)^+$):

$$\begin{aligned}\mathbb{E}_t[(\exp(X_T) - K)^+] &= \int_{\ln K}^{\infty} \exp(x)g(x)\mathrm{d}x - K\int_{\ln K}^{\infty} g(x)\mathrm{d}x \\ &= f(1)\int_{\ln K}^{\infty} \hat{g}(x)\mathrm{d}x - K\int_{\ln K}^{\infty} g(x)\mathrm{d}x. \end{aligned} \tag{D.26}$$

Wendel (1961) provides[b] the following formula for inversion from the characteristic function $f(i\phi)$ to the corresponding distribution function $G(x) = \int_0^x g(s)\mathrm{d}s$:

$$\begin{aligned}G(x) &= \frac{1}{2} - \frac{1}{\pi}\int_0^{\infty} \mathrm{Im}\left[\frac{\exp(-i\phi x)f(i\phi)}{\phi}\right]\mathrm{d}\phi \\ &= \frac{1}{2} - \frac{1}{\pi}\int_0^{\infty} \mathrm{Re}\left[\frac{\exp(-i\phi x)f(i\phi)}{i\phi}\right]\mathrm{d}\phi\end{aligned}$$

where Im and Re denote imaginary and real part respectively. This is used to find:

$$\begin{aligned}\int_{\ln K}^{\infty} g(x)\mathrm{d}x &= 1 - G(\ln K) \\ &= \frac{1}{2} + \frac{1}{\pi}\int_0^{\infty} \mathrm{Re}\left[\frac{K^{-i\phi}f(i\phi)}{i\phi}\right]\mathrm{d}\phi \end{aligned} \tag{D.27}$$

[b] This comes from an original result by Gil-Pelaez (1951), see Lukacs (1970) for more background regarding characteristic functions.

and using that pdf $\hat{g}(x)$ corresponds to characteristic function $\frac{f(i\phi+1)}{f(1)}$ the left hand side of (D.26) is obtained similarly:

$$\int_{\ln K}^{\infty} \hat{g}(x)\mathrm{d}x = \frac{1}{2} + \frac{1}{\pi}\int_0^{\infty} \mathrm{Re}\left[\frac{K^{-i\phi}f(i\phi+1)}{i\phi f(1)}\right]\mathrm{d}\phi. \tag{D.28}$$

Plugging (D.27) and (D.28) into (D.26) gives the pricing formula (4.10), and using that $S_t = e^{-r(T-t)}\mathbb{E}_t^*[S_T]$ gives (4.11).

D.6. Proof of proposition 1

This appendix outlines details of the proof of proposition 1, as given in Christoffersen et al. (2013).

Recall the setting of section 4.1: $(\Omega, \mathcal{F}_T, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P})$ is the filtered probability space that characterises the market, and there exists some risk-neutral equivalent martingale measure $\mathbb{Q}$. It is reminded that $\mathcal{F}_t$-conditional expectation under $\mathbb{P}$ and $\mathbb{Q}$ is denoted as $\mathbb{E}_t[\cdot]$ and $\mathbb{E}_t^*[\cdot]$ respectively.

The pricing kernel (4.4) is rewritten to:

$$\frac{M_t}{M_{t-\Delta}} = \left(\frac{S_t}{S_{t-\Delta}}\right)^{\psi} \exp\left(\delta + \eta h_t + \xi(h_{t+\Delta} - h_t)\right) \tag{D.29}$$

and the HNG-model dynamics (4.1) are rewritten to

$$\begin{aligned} h_{t+\Delta} - h_t &= \omega + (\beta-1)h_t + \alpha\left(z_t - \gamma\sqrt{h_t}\right)^2 \\ \frac{S_t}{S_{t-\Delta}} &= \exp\left(r_{t-\Delta} + \lambda h_t + \sqrt{h_t}z_t\right). \end{aligned}$$

and, dropping subscripts of h_t, z_t, $r_{t-\Delta}$, this is substituted into (D.29) to obtain:

$$\begin{aligned} \frac{M_t}{M_{t-\Delta}} &= \exp\Bigg(\psi(r + \lambda h + \sqrt{h}z) + \delta + \eta h + \\ &\qquad \xi\Big(\omega + (\beta-1)h + \alpha\big(z - \gamma\sqrt{h}\big)^2\Big)\Bigg) \\ &= \exp\big(\omega\xi + \psi r + \delta + (\psi\lambda + \eta + \xi(\beta - 1 + \alpha\gamma^2))h + \\ &\qquad (\psi - 2\xi\alpha\gamma)z\sqrt{h} + \xi\alpha z^2\big). \end{aligned} \tag{D.30}$$

Since the pricing kernel holds for a zero coupon bond (risk free pay-off of 1 at time t) that at time $t-\Delta$ has value $\exp(-r)$, so that in particular $\mathbb{E}_t[M_t] = 1$ and $\mathbb{E}_{t-\Delta}[M_t] = \exp(-r)$, it is found that

$$\mathbb{E}_{t-\Delta}\left[\frac{M_t}{M_{t-\Delta}}\right] = \exp(-r).$$

Equating this to the $t-\Delta$-conditional expectation of (D.30) gives

$$\begin{aligned} \exp(-r) = \exp\big(\omega\xi + \psi r + \delta + \big(\psi\lambda + \eta + \xi(\beta - 1 + \alpha\gamma^2)\big)h\big) \times \\ \mathbb{E}_{t-\Delta}\left[(\psi - 2\xi\alpha\gamma)z\sqrt{h} + \xi\alpha z^2\right] \end{aligned} \tag{D.31}$$

where $\mathcal{F}$-predictability of h was used. A slightly different form of the result for $z \sim \mathcal{N}(0,1)$ and $a < \frac{1}{2}$, (D.21), that was used earlier is invoked:

$$\begin{aligned}\mathbb{E}\left[\exp\left(az^2 + 2abz\right)\right] &= \frac{1}{\exp(ab^2)}\mathbb{E}\left[\exp\left(a(b+z)^2\right)\right] \\ &= \exp\left(-\frac{1}{2}\ln(1-2a) + \frac{ab^2}{1-2a} - ab^2\right) \\ &= \exp\left(-\frac{1}{2}\ln(1-2a) + \frac{2a^2b^2}{1-2a}\right).\end{aligned}$$

In this case:

$$\begin{aligned}a &= \alpha\xi \\ b &= \frac{(\psi - 2\alpha\xi\gamma)\sqrt{h}}{2\alpha\xi}\end{aligned}$$

so that given $\alpha\xi < \frac{1}{2}$ (again easily satisfied for the estimated model parameters) it is obtained that:

$$\begin{aligned}&\mathbb{E}_{t-\Delta}\left[(\psi - 2\xi\alpha\gamma)z\sqrt{h} + \xi\alpha z^2\right] \\ &\qquad = \exp\left(-\frac{1}{2}\ln(1-2\alpha\xi) + \frac{2\alpha^2\xi^2(\psi - 2\alpha\xi\gamma)^2 h}{4\alpha^2\xi^2}\cdot\frac{1}{1-2\alpha\xi}\right) \\ &\qquad = \exp\left(-\frac{1}{2}\ln(1-2\alpha\xi) + \frac{(\psi - 2\alpha\xi\gamma)^2 h}{2-4\alpha\xi}\right).\end{aligned}$$

This is plugged into (D.31). Taking logarithms and rearranging gives:

$$\begin{aligned}\omega\xi + (\psi+1)r + \delta - \frac{1}{2}\ln(1-2\alpha\xi) + \Big(\psi\lambda + \eta + \\ \xi(\beta - 1 + \alpha\gamma^2) + \frac{(\psi - 2\alpha\xi\gamma)^2}{2-4\alpha\xi}\Big)h = 0.\end{aligned} \tag{D.32}$$

Since this holds for any h and $h > 0$, this can be solved to δ and η:

$$\begin{aligned}\delta &= -\omega\xi - (\psi+1)r + \frac{1}{2}\ln(1-2\alpha\xi) \\ \eta &= -\psi\lambda + \xi(1-\beta-\alpha\gamma^2) - \frac{(\psi - 2\alpha\xi\gamma)^2}{2-4\alpha\xi}.\end{aligned} \tag{D.33}$$

Next, using that the pricing kernel holds for a risk-free asset S_t with $S_{t-1} = 1$, so that also $M_{t-1} = 1$, it is observed:

$$\mathbb{E}_{t-\Delta}\left[\frac{S_t}{S_{t-\Delta}}\frac{M_t}{M_{t-\Delta}}\right] = \mathbb{E}_{t-\Delta}\left[M_t S_t\right] = e^r\mathbb{E}_{t-1}[M_t] = e^r e^{-r} = 1$$

and from (D.29) it is found that $\frac{S_t}{S_{t-\Delta}}\frac{M_t}{M_{t-\Delta}}$ is just $M_t/M_{t-\Delta}$ with ψ replaced by $\psi + 1$, so that analogous to (D.32):

$$\begin{aligned}\mathbb{E}\left[\frac{S_t}{S_{t-\Delta}}\frac{M_t}{M_{t-\Delta}}\right] &= \exp\Big(\omega\xi + (\psi+1)r + \delta - \frac{1}{2}\ln(1-2\alpha\xi) + \\ &\qquad \Big((\psi+1)\lambda + \eta + \xi(\beta - 1 + \alpha\gamma^2) + \frac{(\psi + 1 - 2\alpha\xi\gamma)^2}{2-4\alpha\xi}\Big)h\Big) \\ &= 1\end{aligned}$$

and taking logarithms and plugging in (D.33) gives

$$\lambda h + \frac{1 + 2\psi - 4\alpha\xi\gamma}{2 - 4\alpha\xi} h = 0,$$

and dividing by h and solving for ψ gives

$$\begin{aligned}\psi &= -\lambda(1 - 2\alpha\xi) - \frac{1}{2} + 2\alpha\xi\gamma \\ &= -(\lambda + \gamma)(1 - 2\alpha\xi) + \gamma - \frac{1}{2}.\end{aligned}$$

From the $\mathbb{P}$-density function of S_t at $t-1$ (i.e. log-normal PDF with mean $S_{t-\Delta} + r_{t-\Delta} + \lambda h_t$ and variance h_t) and by the fact that the risk-neutral density is equal to the $\mathbb{P}$-density multiplied by the pricing kernel, it can be derived (see Christoffersen et al. (2013)) that under $\mathbb{Q}$ the innovation z_t is normally distributed with mean

$$\mathbb{E}^*_{t-\Delta}[z_t] = -\left(\lambda + \frac{1}{2} + \frac{\alpha\xi}{1 - 2\alpha\xi}\right)\sqrt{h_t}$$

and variance

$$\mathrm{Var}^*_{t-\Delta}[z_t] = \frac{1}{1 - 2\alpha\xi}$$

where Var^*_t denotes $\mathcal{F}_t$-conditional variance under $\mathbb{Q}$. It follows that the following standardisation of z_t is standard-normal under $\mathbb{Q}$:

$$z^*_t = \sqrt{1 - 2\alpha\xi}\left(z_t + \left(\lambda + \frac{1}{2} + \frac{\alpha\xi}{1 - 2\alpha\xi}\right)\sqrt{h_t}\right). \tag{D.34}$$

the risk-neutral process (4.6) follows by plugging in (D.34) into (4.5).

Bibliography

Ait-Sahalia, Y., Karaman, M. & Mancini, L. (2012), 'The Term Structure of Variance Swaps, Risk Premia and the Expectation Hypothesis', *SSRN Electronic Journal* .

Bai, X., Russell, J. R. & Tiao, G. C. (2003), 'Kurtosis of GARCH and stochastic volatility models with non-normal innovations', *Journal of Econometrics* 114(2), 349–360.

Barone-Adesi, G., Engle, R. F. & Mancini, L. (2008), 'A GARCH option pricing model with filtered historical simulation', *Review of Financial Studies* 21(3), 1223–1258.

Bates, D. S. (1996), Testing Option Pricing Models, Technical report, National Bureau of Economic Research. Retrieved April 20, 2020 from: https://ssrn.com/abstract=225194.
URL: *https://ssrn.com/abstract=225194*

Black, F. (1976), Studies in stock price volatility changes, *in* 'In Proceedings of the 1976 American Statistical Association, Business and Economical Statistics Section', pp. 177–181.

Black, F. & Scholes, M. (1973), 'The Pricing of Options and Corporate Liabilities', *Journal of Political Economy* 81(3), 637–654.

Bollerslev, T. (1986), 'Generalized Autoregressive Conditional Heteroskedasticity', *Journal of Econometrics* 31(3), 307–327.

Bollerslev, T. (1987), 'A Conditionally Heteroskedastic Time Series Model for Speculative Prices and Rates of Return', *The Review of Economics and Statistics* 69(3), 542–547.

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

Christoffersen, P., Heston, S. & Jacobs, K. (2013), 'Capturing option anomalies with a variance-dependent pricing kernel', *Review of Financial Studies* 26(8), 1962–2006.

Chu, Shin-herng; Freund, S. (1996), 'Volatility Estimation for Stock index Options: A GARCH Approach', *The Quraterly Reivew of Economics and Finance* 36(4), 431–450.

Cont, R. (2001), 'Empirical properties of asset returns: Stylized facts and statistical issues', *Quantitative Finance* 1(2), 223–236.

Cox, J. C., Ingersoll, J. E. & Ross, S. A. (1985), 'A Theory of the Term Structure of Interest Rates', *Econometrica* .

Damodaran, A. (2002), *Investment Valuation: Second Edition*, Wiley.
URL: *http://people.stern.nyu.edu/adamodar/New_Home_Page/Inv3ed.htm*

Engle, R. F. (1982), 'Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation', *The econometric society* 50(4), 987–1007.

Fama, E. F. (1963), 'Mandelbrot and the Stable Paretian Hypothesis', *The Journal of Business* 36(4), 420–429.

Fama, E. F. (1965), 'The Behavior of Stock-Market Prices', *The Journal of Business* 38(1), 34–105.

Flint, E., Ochse, E. & Polakow, D. (2015), 'Estimating long-term volatility parameters for market-consistent models', *South African Actuarial Journal* 14(1), 19.

Francq, C. & Zakoïan, J. M. (2010), *GARCH Models: Structure, Statistical Inference and Financial Applications*, Wiley.

Gil-Pelaez, J. (1951), 'Note on the Inversion Theorem', *Biometrika Trust* 38(3), 481–482.
URL: *https://www.jstor.org/stable/2332598%0D*

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

Heston, S. L. & Nandi, S. (2000), 'A closed-form GARCH option valuation model', *Review of Financial Studies* 13(3), 585–625.

Hsieh, P. L., Zhang, Q. Q. & Wang, Y. (2018), 'Jump risk and option liquidity in an incomplete market', *Journal of Futures Markets* 38(11), 1334–1369.

Hull, J. & White, A. (1987), 'The Pricing of Options on Assets with Stochastic Volatilities', *The Journal of Finance* 42(2), 281.

Kozhan, R., Neuberger, A. & Schneider, P. (2013), 'The skew risk premium in the equity index market', *Review of Financial Studies* 26(9), 2174–2203.

Kumar, R. & Dhankar, R. S. (2010), 'Empirical analysis of conditional heteroskedasticity in time series of stock returns and asymmetric effect on volatility', *Global Business Review* 11(1), 21–33.

Lassance, N. & Vrins, F. (2018), 'A comparison of pricing and hedging performances of equity derivatives models', *Applied Economics* 50(10), 1122–1137.
URL: *https://doi.org/10.1080/00036846.2017.1352080*

Leisen, D. P. (2017), 'The shape of small sample biases in pricing kernel estimations', *Quantitative Finance* 17(6), 943–958.

Lukacs, E. (1970), 'Characteristic Functions.', *The Incorporated Statistician* .

Mandelbrot, B. (1963), 'The Variation of Certain Speculative Prices', *The Journal of Business* .

Merton, R. (1973), 'Theory of Rational Option Pricing', *Bell Journal of Economics* 4(1), 141–183.

Miron, D. & Tudor, C. (2010), 'Asymmetric conditional volatility models: Empirical estimation and comparison of forecasting accuracy', *Romanian Journal of Economic Forecasting* 13(3).

Moyaert, T. & Petitjean, M. (2011), 'The performance of popular stochastic volatility option pricing models during the subprime crisis', *Applied Financial Economics* 21(14), 1059–1068.

Nelson, D. B. (1990), 'ARCH models as diffusion approximations', *Journal of Econometrics* 45(1-2), 7–38.

Nelson, D. B., Foster, D. P., Models, A., Nelson, B. Y. D. B. & Foster, D. P. (1994), 'Asymptotic Filtering Theory for Univariate Arch Models', *Econometrica* 62(1), 1–41.
URL: *http://www.jstor.org/stable/2951474*

Ortec Finance (2018), Methodology of Risk Neutral Scenarios, Technical report, Risk-Neutral department, Ortec Finance.

Rubinstein, M. (1985), 'Nonparametric Tests of Alternative Option Pricing Models Using All Reported Trades and Quotes on the 30 Most Active CBOE Option Classes from August 23, 1976 through August 31, 1978', *The Journal of Finance* 40(2), 455–480.

Shephard, N. & Andersen, T. G. (2009), Stochastic Volatility: Origins and Overview, *in* 'Handbook of Financial Time Series', Elsevier.

Shu, J. & Zhang, J. E. (2003), 'The relationship between implied and realized volatility of S&P 500 index', *Wilmott* 2003(1), 83–91.

Steehouwer, H. (2019), Ortec Finance Scenario approach, Technical report, Ortec Finance Scenario Department Paper.

Su, Y. C., Chen, M. D. & Huang, H. C. (2010), 'An application of closed-form GARCH option-pricing model on FTSE 100 option and volatility', *Applied Financial Economics* 20(11), 899–910.

Van Dijk, M., de Graaf, C. & Oosterlee, C. (2018), 'Between P and Q: The PQ Measure for Pricing', *Journal of Risk and Financial Management* 11(4), 67.

Wang, T., Shen, Y., Jiang, Y. & Huang, Z. (2017), 'Pricing the CBOE VIX Futures with the Heston–Nandi GARCH Model', *Journal of Futures Markets* 37, 641–659.

Wendel, J. G. (1961), 'The Non-Absolute Convergence of Gil-Pelaez' Inversion Integral', *The Annals of Mathematical Statistics* 32(1), 338–339.