Detecting Money Market Bubbles
Jan Baldeaux\textsuperscript{1}, Katja Ignatieva\textsuperscript{2}, Eckhard Platen\textsuperscript{3}

February 18, 2014

Abstract. Money bubbles, as formally introduced by Heston, Loewenstein and Willard in 2007, preclude the existence of a risk-neutral pricing measure. This is due to the fact that the existence of a money market bubble implies the existence of a self-financing trading strategy that replicates the savings account’s value at a fixed future date at a cheaper cost than the current value of the savings account. From the point of view of derivative pricing, it is hence crucial to understand if markets exhibit money market bubbles. The benchmark approach is formulated under the real world probability measure and does not require the existence of a risk neutral probability measure. It hence emerges as the appropriate framework to study the potential existence of money market bubbles. Using a range of stochastic volatility models well-known in the finance literature, we study the existence of money market bubbles in the US economy, and find that for all models the US market exhibits a money market bubble. This conclusion suggests that for derivative pricing and hedging care should be taken when making assumptions pertaining to the existence of a risk-neutral measure. Less expensive hedge portfolios may exist for a wide range of derivatives.

\textit{JEL Classification: C6, C63, G1, G13}

\textit{Key words and phrases: Bubbles; Money Market Bubbles; Strict Local Martingales; Markov Chain Monte Carlo; Stochastic Volatility Models; Benchmark Approach}

\textsuperscript{1}University of Technology Sydney, Finance Discipline Group, PO Box 123, Broadway, NSW, 2007, Australia. Email: jan.baldeaux@uts.edu.au

\textsuperscript{2}University of New South Wales Sydney, Risk & Actuarial School, Australian School of Business, Kensington, NSW, 2052, Australia. Email: k.ignatieva@unsw.edu.au

\textsuperscript{3}University of Technology Sydney, Finance Discipline Group and School of Mathematical Sciences, PO Box 123, Broadway, NSW-2007, Australia. Email: eckhard.platen@uts.edu.au
1 Introduction

In this paper, we empirically investigate the potential existence of money market bubbles, as formally introduced in the literature in [27]. Already in [54], the potential existence of this type of arbitrage was theoretically pointed out. Money market bubbles form part of a growing literature on bubbles, which are studied in both, the economics and the finance literature, see e.g. [14], [23], [24], [44], [33], [27], [28], [29], [32], [34], [35], and in particular the recent article [50] for an overview of the literature. In [27], an asset price bubble is defined as an asset with a nonnegative price, which can be replicated at a fixed future time using a self-financing trading strategy whose setup cost is lower than the current value of the asset. As in [27], we place ourselves in a two asset economy, and assume the existence of a well-diversified index and a money market account. A money market bubble exists if the savings account can be replicated using the well-diversified index at a cost that is cheaper than the value of the savings account. As established in [27], a necessary and sufficient condition for the existence of a bubble is the failure of the existence of an equivalent martingale measure (EMM), i.e. the putative Radon-Nikodym derivative of the EMM is a strict local martingale. So far though, only the paper [27] and [29] have dealt with money market bubbles. Furthermore, we point out that under the benchmark approach, [49], money market bubbles naturally occur, as for example under the Minimal Market Model (MMM), see also [49] and [54]. Should the candidate model rule out money market bubbles, one can investigate if the discounted well-diversified index follows a strict local martingale under a respective EMM. If this is the case, one has detected an asset-pricing bubble. The latter have received much attention in the
Figure 1.1: Empirical Radon-Nikodym derivative for the US economy.

literature, see e.g. [32], [34], [35], and in particular the recent article [50]. The conditions for the existence of asset pricing bubbles are by now well understood, an asset pricing bubble exists if and only if the discounted index follows a strict local martingale under the EMM, and asset pricing bubbles exhibit some interesting financial consequences, such as failure of put-call parity, non-convexity of option prices with respect to the underlying for convex options with convex payoff functions, to name but a few.

This paper examines the fundamental question of the existence of an EMM when aiming for a candidate model. A motivation for this type of examination is the plot of the US savings account in units of the total S&P 500 return, shown in Figure 1.1, which displays the empirical Radon-Nikodym derivative for the US market, for a complete index model that interprets the S&P 500 as numéraire portfolio, as will be explained later.
The empirical Radon-Nikodym derivative in Figure 1.1 demonstrates a systematic downward trend which is indicative of a nonnegative strict local martingale, which is a supermartingale, and hence the presence of a money market bubble. This is opposed to a martingale, where the present value is the best forecast of any future values. We conclude that Figure 1.1 serves as a motivation for a thorough investigation of the potential existence of money market bubbles.

To investigate the existence of an EMM, we require an approach more general than the classical non-arbitrage approach with its risk neutral pricing. In particular, we require an approach which is formulated under the real world probability measure and more classical than the general approach. The benchmark approach, [49], offers this flexibility and allows us to rigorously examine the potential existence of an EMM. At the heart of the benchmark approach sits the numéraire portfolio (NP), see [43], which when used as benchmark, makes all nonnegative benchmarked portfolios supermartingales. For most models it is identical to the growth optimal portfolio (GOP), which is mathematically defined as the portfolio which maximizes logarithmic utility from terminal wealth. It plays a fundamental role in derivative pricing, as it serves as numéraire when pricing contingent claims under the real world probability measure. We alert the reader to the fact that when employing the benchmark approach for derivative pricing, one finds the minimal possible price by employing the real world pricing formula, [49], which computes the expectation under the real world measure using the NP as numéraire. In terms of money market bubbles, we recall that a money market bubble exists if the savings account can be replicated using the NP at a setup cost cheaper than the current value of the savings account. The benchmark approach
allows us to empirically answer this question. More precisely, this paper answers the following questions:

- Does a money market bubble exist in the US market?

- Which is a tractable, parsimonious, yet sufficiently accurate model for the discounted S&P 500?

We answer the first question by connecting two significant streams in the finance literature, namely the literature concerned with the estimation of stochastic volatility models and the literature on the existence of bubbles. Using Markov Chain Monte Carlo (MCMC) methods, see [13], [20], and also the references therein, on parameterized stochastic volatility models describing the S&P 500 index. We adopt this approach and fit various stochastic volatility models that are well established in the finance literature to the S&P 500 index, including the Heston, the 3/2, and the continuous time Garch model. We refer to [41] for an overview of stochastic volatility models, including the latter ones. We remark at this point that there are some criticisms pertaining to some of these models, as for example discussed in [12] and [15]. For all stochastic volatility models considered in this paper, we can present necessary and sufficient conditions for the existence of a money market bubble. Under the Heston model, we demonstrate that money market bubbles cannot exist, it mainly serves as a reference model. However, for the remaining models, we find that for some parameter regimes, money market bubbles exist. In fact, we find that when fitting these models, they all produce money market bubbles. This suggests that assumptions pertaining to the existence of an EMM should be made with care. Secondly, we find that the $3/2$
stochastic volatility model, which is a special case of the model studied in [5], provides the best fit to empirical data.

This paper can be seen as a companion paper to [27], [29], [32], and [6]. It is the first paper that tests for the existence of money market bubbles, which have only been discussed in [27], [29], in the same way that [32] was the first paper to investigate empirically the existence of stock price bubbles. Finally, we remark that in [6], local volatility models were fitted to well-diversified indices using techniques which were extended in [32], and used in the latter paper to examine the existence of stock price bubbles. We point out that the results presented in [6] support our findings which suggest the existence of money market bubbles in the markets under consideration.

We conclude the introduction with a brief discussion of the plausibility of money market bubbles. The presence of money market bubbles violates the "no free lunch with vanishing risk" (NFLVR) condition, see [16], [17], which is the no-arbitrage condition, which underpins classical risk neutral pricing. But the models we present still satisfy the more general "no unbounded profit with bounded risk" (NUPBR) condition, see [38], which is sufficient for utility maximization in a general semimartingale framework. They also fit under the benchmark approach, which only assumes the existence of the numéraire portfolio and automatically excludes, so called, strong arbitrage, that would create strictly positive wealth from a nonnegative portfolio with zero initial capital, see [49]. From the point of view of economic intuition though, a reader acquainted with derivative pricing based on the absence of classical arbitrage opportunities might challenge the existence of money market bubbles: Shorting the asset and replicating it using
the index would then produce a certain positive cash-flow either at the time the arbitrage is set up or equivalently at the time when the arbitrage strategy matures. One could argue that arbitrageurs would enter into the arbitrage strategy until prices converged, i.e. the bubble disappeared. The key problem with the above argument is that it is essentially a "one-period model" argument: One only considers the portfolio value at two points in time, when the arbitrage is set up, and when the arbitrage matures. However, one ignores the portfolio value evolution between these two time points, which in fact could become negative with strictly positive probability. This has important (corporate finance) implications, as liabilities must be secured by posting collateral. This is explored in detail in [42], where a model is presented in which an EMM does not exist, and in which arbitrages are considered as risky investment, since they can also produce negative returns. It is shown that it can be optimal for the investor to underinvest in the arbitrage opportunity and not to take the largest arbitrage position allowed by the collateral constraints. Clearly, if the investor only takes a limited position in the arbitrage position or avoids taking a position in the arbitrage opportunity, then there is no reason for the bubble to disappear or not to widen.

The remainder of the paper is structured as follows: in Section 2, we introduce the modeling framework which will be used to detect money market bubbles, in Section 3 we introduce the statistical methodology used to detect bubbles, and in Section 4 we present the empirical results. Section 5 interprets the results and Section 6 concludes the paper.
2 Money Market Bubbles

In this section, we introduce the framework which will allow us to empirically test the existence of money market bubbles. We place ourselves in the setup presented in [27], Section 2, and formulate our model on a filtered probability space \((\Omega, A, A, P)\), where \(P\) is the real world probability measure and \(A = (A_t)_{t \geq 0}\) is a filtration satisfying the usual conditions. We assume that the squared volatility \(V_t\) of the NP, which is a state variable, is the unique strong solution of the stochastic differential equation (SDE)

\[
dV_t = a(V_t)dt + b(V_t)dW^1_t, \tag{2.1}
\]

where \(V_0\) is strictly positive. The process \(W^1_t\) is a Brownian motion under the empirical probability measure \(P\). In the following subsections we introduce specific models for \(V_t\), specifying \(a(\cdot)\) and \(b(\cdot)\). In particular, we assume that the functions \(a(\cdot)\) and \(b(\cdot)\) employed in this paper are continuous functions on \((0, \infty)\) with \(b^2(\cdot) > 0\). Secondly, we present a model for the discounted numéraire portfolio (NP). In particular, we assume that the discounted NP satisfies the SDE

\[
dS_t = S_t \left( V_t dt + \sqrt{V_t} dW^2_t \right), \tag{2.2}
\]

where \(\sqrt{V_t}\) denotes the volatility and the market price of risk, which we specify further below, and \(W^2_t\) is a Brownian motion. Finally, we introduce the money market account,

\[
 dB_t = r_t B_t dt, \tag{2.3}
\]

where \(B_0 = 1\). Finally we assume that the Brownian motions \(W^1_t\) and \(W^2_t\) are correlated, their constant correlation being \(\rho\). We now provide a definition for a money market bubble, which is taken from [27], see Definition 2.1.
**Definition 2.1.** The money market account has a bubble, if there is a self-financing portfolio with pathwise nonnegative wealth that costs less than the money market and replicates the money market at a fixed future date. The value of the money market bubble is the difference between the money market's price and the lowest cost replicating strategy.

Under the benchmark approach the NP is employed as benchmark. Next, we introduce the benchmarked money market account $\hat{B} = \{\hat{B}_t = \frac{1}{S_t}, t \geq 0\}$, which is simply the inverse of the discounted GOP and plays a central role in the detection of bubbles. The dynamics of $\hat{B}$ are given by

$$d\hat{B}_t = -\sqrt{V_t}\hat{B}_tdW_t, \quad (2.4)$$

to obtain

$$\hat{B}_t = \hat{B}_0 \exp\left(-\int_0^t \sqrt{V_s}dW_s - \frac{1}{2} \int_0^t V_s ds\right).$$

We now introduce the Radon-Nikodym derivative for the putative equivalent martingale measure (EMM), which is given by $\xi_t = \frac{\hat{B}_t}{\hat{B}_0}, t \geq 0$. We now recall Proposition 2.1 from [27].

**Proposition 2.1.** The process $\xi = \{\xi_t, t \geq 0\}$ is a strict local martingale if and only if the money market account has a bubble.

As shown in [27], the process $\xi$ is a strict local martingale, if and only if the process $\tilde{V}$ does not reach zero or infinity in finite time, where $\tilde{V}$ is given by the SDE

$$d\tilde{V}_t = (a(\tilde{V}_t) - \rho \sqrt{\tilde{V}_t}b(\tilde{V}_t))dt + b(\tilde{V}_t)dB_t, \quad (2.5)$$

where $\tilde{V}_0 = V_0$. We formalize this in the following Corollary, which reflects the Condition 2' from [27].
Corollary 2.1. The money market account exhibits a bubble if and only if the solution of (2.5) explodes or hits 0 in finite time.

There are necessary and sufficient conditions which allow one to check if a particular one-dimensional diffusion, such as the one in equation (2.5), reaches 0 or infinity. This includes the Fellers test, see e.g. Theorem 5.29 (Section 5.5) in [39], which states.... This argument appears in [52], but has subsequently been used to study whether local martingales are strict or not, see e.g. [49] and the references therein; and also to study moment explosions, see e.g. [3].

We now discuss the four stochastic volatility models we wish to discuss in the paper. All fall into the above framework.

2.1 Heston Model

For the Heston model, see [26], we set

\[ a(x) = \kappa(\theta - x), \quad b(x) = \sigma \sqrt{x}, \]

in equation (2.1), where \( \kappa > 0, \theta > 0, \) and \( \sigma > 0. \) We assume that the Feller condition is in force, i.e. that \( 2\kappa\theta > \sigma^2. \) We then have the following result, which follows immediately from Corollary 2.1.

**Theorem 2.1.** Assume that the dynamics of \( \hat{B} \) are given by the Heston model. Then \( \hat{B} \) is a martingale and the money market account does not exhibit a bubble.

**Proof.** The result can be directly established using the argument from [3], which is given in Lemma 2.3. \( \square \)
We remark that based on Figure 1.1, the Heston model does not seem to provide a suitable model for the discounted S&P 500, since the $\hat{B}_t$ evolves more likely as a strict local martingale than a martingale. Nevertheless, the Heston model will serve us for comparisons. Furthermore, we find it interesting to investigate whether empirically the Feller condition is satisfied or not.

### 2.2 3/2 Nonlinearly Mean Reverting Model

To formulate the nonlinearly mean reverting 3/2 model (3/2N), introduced in [1] and [47], and further developed in [41], we set

$$a(x) = \kappa x (\theta - x), \quad b(x) = \sigma x^{3/2},$$

where $\kappa > 0$, $\theta > 0$, and $\sigma > 0$. Under the 3/2N model, for certain parameter combinations the process $\hat{B}$ is a martingale, but otherwise a strict local martingale.

**Theorem 2.2.** Assume that the dynamics of $\hat{B}$ are given by the 3/2N model. Then $\hat{B}$ is a martingale, if and only if

$$\frac{\kappa}{\sigma} + \frac{\sigma}{2} \geq -\rho.$$

**Proof.** This result was first proven in [5] to which we refer.

In Section 3, we will estimate the parameters of the 3/2N model, to determine whether $\hat{B}$ follows potentially a strict local martingale.
2.3 Linearly Mean Reverting 3/2 Model

To formulate the linearly mean reverting 3/2L model, see [37], we set

\[ a(x) = \kappa(\theta - x), \quad b(x) = \sigma x^{3/2}, \]

where \( \kappa > 0, \theta > 0, \) and \( \sigma > 0. \) For this model, we obtain that for certain parameter combinations the process \( \hat{B} \) is a martingale, but otherwise a strict local martingale.

**Theorem 2.3.** Assume that the dynamics of \( \hat{B} \) are given by the linearly mean reverting 3/2 model. Then \( \hat{B} \) is a martingale, if and only if

\[ \frac{\sigma}{2} \geq -\rho. \]

**Proof.** The proof follows from Proposition 2.5 in [3]. Since we consider the benchmarked savings account, we replace \( \rho \) in Proposition 2.5 by \( -\rho. \)

2.4 GARCH Model

Similar to the conditions limit of the GARCH model, see [48], we consider also the GARCH model, which was e.g. studied in [25]. It is characterized by the choice

\[ a(x) = \kappa(\theta - x), \quad b(x) = \sigma x, \]

where \( \kappa > 0, \theta > 0, \) and \( \sigma > 0. \) We obtain that for certain parameter combinations of this model that the process \( \hat{B} \) is a martingale, but otherwise a strict local martingale.
Theorem 2.4. Assume that the dynamics of \( \hat{B} \) are given by the GARCH model. Then \( \hat{B} \) is a martingale, if and only if 

\[ \rho > 0. \]

Proof. The proof follows from Proposition 2.5 in [3]. Since we consider the benchmarked savings account, we replace \( \rho \) in Proposition 2.5 by \( -\rho \).

Note that the modeling framework introduced above is flexible enough to be extended to multi-factor stochastic volatility models and models which include jumps. However, for the purpose of this paper we concentrate on the one-factor stochastic volatility models. In the next section, we discuss the statistical methodology which will be used to estimate the models presented in this section. Having estimated the models, we can use Theorems 2.2, 2.3, and 2.4, respectively, to determine if a money market bubble is present or not assuming the respective model.

3 Markov-Chain Monte-Carlo Estimation

We now discuss the statistical methodology used in this paper, which is based on [13], [20]. Firstly, we define log-returns on the benchmarked savings account by considering the increments of \( Y_t = \log(\hat{B}_t) \), which satisfies the SDE

\[ dY_t = -\frac{1}{2}V_t dt - \sqrt{V_t} dW_t. \]

By employing a time discretization with \( t_k = k\Delta, \Delta > 0, k \in \{0,1,\ldots\} \) and the Euler discretization scheme, we consider \( R_{t_k+1} = Y_{t_{k+1}} - Y_{t_k} \), so that

\[ R_{t_k+1} = -\frac{1}{2} \int_{t_k}^{t_{k+1}} V_s ds - \int_{t_k}^{t_{k+1}} \sqrt{V_s} dW_s. \]
\[ -\frac{1}{2} V_{t+k} \Delta - \sqrt{V_{t+k}} (W_{t+k+1} - W_{t+k}) \]

\[ = -\frac{1}{2} V_{t+k} \Delta - \sqrt{V_{t+k}} \varepsilon_{t+k+1}^y, \]

where \( \varepsilon_{t+k+1}^y = W_{t+k+1} - W_{t+k} \). For the squared volatility processes we apply similarly the Euler scheme and specification,

\[
V_{t+k+1} = V_{t+k} + \int_{t+k}^{t+k+1} a(V_s) ds + \int_{t+k}^{t+k+1} b(V_s) dW_s^1
\]

\[ \approx V_{t+k} + a(V_{t+k}) \Delta + b(V_{t+k})(W_{t+k+1}^1 - W_{t+k}^1). \]

We note that all models discussed in this paper can be represented as follows,

\[ V_{t+k+1} \approx V_{t+k} + \kappa (V_{t+k})^a (\theta - V_{t+k}) \Delta + \sigma V_{t+k}^b \varepsilon_{t+k+1}^v \]

for \( a \in \{0, 1\} \) and \( b \in \left\{ \frac{1}{2}, 1, \frac{3}{2} \right\} \) where \( \varepsilon_{t+k+1}^v = W_{t+k+1}^1 - W_{t+k}^1 \). In particular, we have

\[
V_{t+k+1} = V_{t+k} + \kappa (\theta - V_{t+k}) \Delta + \sigma V_{t+k}^{3/2} \varepsilon_{t+k+1}^v \quad (a = 1, b = 3/2: \ 3/2N)
\]

\[
V_{t+k+1} = V_{t+k} + \kappa (\theta - V_{t+k}) \Delta + \sigma V_{t+k}^{3/2} \varepsilon_{t+k+1}^v \quad (a = 1, b = 3/2: \ 3/2L)
\]

\[
V_{t+k+1} = V_{t+k} + \kappa (\theta - V_{t+k}) \Delta + \sigma V_{t+k} \varepsilon_{t+k+1}^v \quad (a = 0, b = 1: \ \text{GARCH})
\]

The underlying problem setup involves the estimation of the parameter vector \( \Theta = (\theta, \kappa, \sigma, \rho)^T \) as well as the volatility vector \( V \), which is a state variable. In a Bayesian context each of these unobserved variables is treated as a parameter, that is, to be estimated. This leads to a high-dimensional posterior distribution, which is not known. In order to compute the moments of the posterior, we would have to compute a high dimensional integral, which is not feasible. Therefore, we rely on the Markov-Chain Monte-Carlo (MCMC) method to compute the moments of the parameter values and latent variables conditional on the observed
data. These moments are used as a point estimator for the parameters and the latent variables. Bayesian MCMC methods for stochastic volatility models have been developed in [31]. MCMC generates samples from a given target distribution, in our case \( p(\Theta, V | R) \) - the joint distribution of the parameter vector \( \Theta \) and the state variable vector \( V \), given the observed vector of returns \( R \). The basis for MCMC estimation is provided by the Hammersley-Clifford theorem, see ..., which claims that under certain regularity conditions, the joint posterior distribution can be completely characterized by the complete conditional distributions.

In order to compute the posterior distribution note that the joint distribution of \((R_{tk}, V_{tk})\) follows approximately a bivariate normal distribution with the following parameters:

\[
\begin{align*}
\mu_{tk-1} &= \begin{pmatrix}
-\frac{1}{2}V_{tk-1}\Delta \\
V_{tk-1} + \kappa V_{tk-1}^a (\theta - V_{tk-1})\Delta 
\end{pmatrix} \\
\Sigma_{tk-1} &= \begin{pmatrix}
V_{tk-1} & -\rho \sigma V_{tk-1}^{b+0.5} \\
-\rho \sigma V_{tk-1}^{b+0.5} & \sigma^2 V_{tk-1}^{2b}
\end{pmatrix},
\end{align*}
\tag{3.6}
\]

\( t = k\Delta, \ k \in \{1, 2, ...\} \) Given the normality of the joint distribution we can determine the conditional distribution of returns \( p(R_{tk} | V_{tk}, V_{tk-1}, \Theta) \). It follows a normal distribution with parameters

\[
\begin{align*}
\mu_{R_{tk}|V_{tk}} &= -\frac{1}{2}V_{tk-1}\Delta - \rho \sigma^{-1}V_{tk-1}^{0.5-b}(V_{tk} - V_{tk-1}) - \kappa V_{tk-1}^a (\theta - V_{tk-1})\Delta \\
\sigma^2_{R_{tk}|V_{tk}} &= V_{tk-1}(1 - \rho^2).
\end{align*}
\tag{3.7}
\]

On the other hand, the conditional distribution for \( p(V_{t|R_{t}, V_{t-1}, \Theta}) \) we obtain a normal distribution with

\[
\begin{align*}
\mu_{V_{tk}|R_{tk}} &= V_{tk-1} + \kappa (V_{tk-1})^a (\theta - V_{tk-1})\Delta - \rho \sigma V_{tk-1}^{b-0.5} (R_{tk} + \frac{1}{2}V_{tk-1})\Delta \\
\sigma^2_{V_{tk}|R_{tk}} &= \sigma^2 V_{tk-1}^{2b}(1 - \rho^2).
\end{align*}
\]
In general, the posterior is given by the following expression:

\[
p(\Theta, V|R) \propto p(R|\Theta, V)p(\Theta, V) = p(R|\Theta)p(V|\Theta)p(\Theta),
\]  

(3.8)

where \(p(R|\Theta, V)\) denotes the likelihood and \(p(\Theta, V)\) is the prior distribution.

In (3.8) we can write \(p(R|\Theta, V) = \prod_{k=1}^{n} p(R_{tk}|V_{tk}, V_{tk-1}, \Theta)\) by conditional independence and \(p(V|\Theta) \propto \prod_{k=1}^{n} p(V_{tk}|V_{tk-1}, \Theta)\) by the Markov property. Using the fact that

\[
p(R|\Theta, V)p(V|\Theta) \propto \prod_{k=1}^{n} p(R_{tk}|V_{tk}, V_{tk-1}, \Theta)p(V_{tk}|V_{tk-1}, \Theta)
\]

\[= \prod_{k=1}^{n} p(R_{tk}, V_{tk}|V_{tk-1}, \Theta),\]

(3.9)

we can rewrite (3.8) as follows:

\[
p(\Theta, V|R) \propto \prod_{k=1}^{n} p(R_{tk}, V_{tk}|V_{tk-1}, \Theta)p(\Theta).
\]  

(3.10)

To update estimated parameter values in each iteration, the MCMC algorithm draws from its posterior distribution conditional on the current values of all other parameters and state variables. Therefore, in order to reduce the influence of the starting point and to assure that stationarity is achieved, the general approach is to discard a burn-in period of the first \(h\) iterations. The iterations after the burn-in period provide a representative sample from the joint posterior, and averaging over the non-discarded iterations provides an estimate for posterior means of parameters and latent variables.

Sampling from the conditional posterior can be implemented by either using a Gibbs sampler introduced by [21] or the Metropolis-Hasting algorithm, see [45].
A Gibbs sampler is applied if the complete conditional distribution to sample from is known. The MCMC algorithm with the Gibbs step samples iteratively drawing from the following conditional posteriors, see [22]:

- **Parameters**: \( p(\Theta_i|\Theta_{-i}, V, R) \), \( i = 1, \ldots, M \)
- **Volatility**: \( p(V_t|\Theta, V_{t+1}, V_{t-1}, R_t, R_{t+1}) \), \( k = 1, \ldots, n \)

where \( M \) denotes the number of parameters, \( n \) is the number of observations and \( \Theta_{-i} \) is the parameter vector \( \Theta \) without the \( i \)-th element.

To start the procedure, we have to specify the prior distributions. When possible we assume so-called conjugate priors which after multiplying with the likelihood lead to a posterior distribution belonging to the same family of distributions as the prior itself. Wherever possible, we choose standard conjugate priors, which allow to draw from the conditional posteriors directly. For the model parameters we specify conjugate priors in Appendix 7.

If some conditional distributions cannot be sampled directly, as in the case with the squared volatility or variance of the log-returns, where the complete conditional distribution is not easily recognizable, we apply the Metropolis-Hastings algorithm. For details on the Metropolis-Hastings algorithm we refer to [36]. We have for the full joint posterior the formula

\[
p(V|\Theta, R) \propto \prod_{k=1}^{n} p(R_{t_k}, V_{t_k}|V_{t_k-1}, \Theta)p(\Theta) \]

Applying the above expression to \( V_{t_k} \) and removing all terms that do not include \( V_{t_k} \) directly, leaves

\[
p(V_{t_k}|\Theta, R) \propto p(R_{t_k}|V_{t_k}, V_{t_k-1}, \Theta)p(V_{t_k}|V_{t_k-1}, \Theta)p(R_{t_k+1}|V_{t_k+1}, V_{t_k}, \Theta)p(V_{t_k+1}|V_{t_k}, \Theta).
\]
Note, that since

\[
p(R_{t_k}, V_{t_k} | V_{t_k-1}, \Theta) = p(R_{t_k} | V_{t_k}, V_{t_k-1}, \Theta)p(V_{t_k} | V_{t_k-1}, \Theta)
= p(V_{t_k} | R_{t_k}, V_{t_k-1}, \Theta)p(R_{t_k} | V_{t_k-1}, \Theta),
\]

we can alternatively use the following expression for the full joint posterior

\[
p(V | \Theta, R) \propto \prod_{k=1}^n p(R_{t_k}, V_{t_k} | V_{t_k-1}, \Theta)p(\Theta)
= \prod_{k=1}^n p(V_{t_k} | R_{t_k}, V_{t_k-1}, \Theta)p(R_{t_k} | V_{t_k-1}, \Theta)p(\Theta)
\propto \prod_{k=1}^n p(V_{t_k} | R_{t_k}, V_{t_k-1}, \Theta)p(R_{t_k} | V_{t_k-1}, \Theta).
\]

Thus, the complete conditional distribution for \( V_{t_k} \) (after removing all terms which do not depend on \( V_{t_k} \)) reduces to

\[
p(V_{t_k} | \Theta, R) \propto p(V_{t_{k+1}} | R_{t_{k+1}}, V_{t_k}, \Theta)p(R_{t_{k+1}} | V_{t_k}, \Theta)p(V_{t_k} | R_{t_k}, V_{t_k-1}, \Theta)
= p(R_{t_{k+1}} | V_{t_{k+1}}, \Theta)p(V_{t_k} | R_{t_k}, V_{t_k-1}, \Theta).
\]

(3.12)

In the last expression, \( p(R_{t_{k+1}} | V_{t_{k+1}}, \Theta) \) follows a bivariate normal distribution and \( p(V_{t_k} | R_{t_k}, V_{t_k-1}, \Theta) \) follows a univariate normal distribution. The Metropolis-Hasting step proposes a new variance \( V_{t_k}^{(j)} \) in the \( j^{th} \) draw by drawing from \( p(V_{t_k} | R_{t_k}, V_{t_k-1}, \Theta) \) and accepting that draw with probability

\[
\min \left\{ \frac{p(R_{t_{k+1}} | V_{t_{k+1}}, V_{t_k}^{(j)}, \Theta)}{p(R_{t_{k+1}} | V_{t_{k+1}}, V_{t_k}^{(j-1)}, \Theta)}, 1 \right\}.
\]

(3.13)

4 Model Testing

Recall that in addition to detecting money market bubbles, this paper also aims to determine which model fits best the discounted S&P 500 data. Generally, it
is not easy to compare different model specifications since a variety of decision criteria exists. This section aims to present several decision criteria which quantify the model performance or fit to the data. The first procedure to investigate the model performance is the visual inspection achieved by comparing quantile-to-quantile (QQ) plots. Since residuals from the return and the variance equations are assumed to follow normal distributions, one can contrast the quantiles of these error distributions against normal quantiles in the QQ plot. The degree of model misspecification (non-normality of residuals) is defined by how far the estimated residuals in log-return and the variance equation deviate from the 45 degrees line. Secondly, we will use the deviance information criterion (DIC) proposed in [53], which allows us to compare non-nested models. Finally, we compare the estimated volatility path with the path of the Chicago Board Options Exchange Market Volatility Index (VIX)\(^4\), as it has been the best measure capturing the market perspective on the implied volatility of the S&P 500 index options. In order to allow the comparison with the VIX, we will use the S&P 500 index as our proxy for the NP in Section 5. The following subsection deals with the DIC in more details.

### 4.1 Deviance Information Criterion

In a non-Bayesian setting deviance is used as a quantity which estimates the number of the parameters in the underlying model: it refers to the difference in log-likelihoods between the fitted and the saturated model, that is the one which yields a perfect fit of the data. Model comparison can be performed based on

the fit, which is the deviance statistic and the complexity, which is the number of free parameters in the model. In analogy to the well-known criteria like the AIC introduced in [2] or BIC constructed in [51], the authors of [18], [53] and [8] have developed the deviance information criterion (DIC) as a Bayesian model choice criterion. DIC solves the problem of comparing complex hierarchial models (among others Bayesian models) when the number of parameters is not clearly defined. DIC consists of two components: a term $\bar{D}$ that measures the goodness of fit and a penalty term $p_D$ accounting for the model complexity:

$$DIC = \bar{D} + p_D.$$

The first term can be calculated as follows:

$$\bar{D} = E_{\Theta|R}\{D(\Theta)\} = E_{\Theta|R}\{-2 \log f(R|\Theta)\},$$

where $R$ denotes the returns data and $\Theta$ is a vector of parameters. The better the model fits data, the larger is the likelihood, i.e. smaller values of $\bar{D}$ will indicate a better model fit. In fact, since DIC already includes a penalty term $p_D$, it could be better thought of as a measure of ‘model adequacy’ rather than a measure of fit, although these terms can be used interchangeably.

The second component measures the complexity of the model by the effective number of parameters:

$$p_D = \bar{D} - D(\Theta) = E_{\Theta|R}\{D(\Theta)\} - D\{E_{\Theta|R}(\Theta)\}$$

$$= E_{\Theta|R}\{-2 \log f(R|\Theta)\} + 2 \log f(R|\bar{\Theta}).$$

Clearly, since $p_D$ is considered to be the posterior mean of the deviance (average of log-likelihood ratios) minus the deviance evaluated at the posterior mean
(likelihood evaluated at average), it can be used to quantify the number of free parameters in the model. Furthermore, defining \( -2 \log f(R|\Theta) \) to be the residual information in the data \( R \) conditional on \( \Theta \), and interpreting it as a logarithmic penalty, or uncertainty, see [40], [9], \( p_D \) can be regarded as the expected excess value of the true over the estimated residual information in data \( R \) conditional on \( \Theta \), and thus, can be thought of as the expected reduction in uncertainty.

From (4.16) we obtain: \( \bar{D} = D(\bar{\Theta}) + p_D \), and thus, DIC can be rewritten as the estimate of the fit plus twice the number of effective parameters (measure of complexity):

\[
\text{DIC} = D(\bar{\Theta}) + 2p_D.
\]

(4.17)

5 Empirical Analysis

In this section we analyze the performance of the different models using time series of daily long-returns for the benchmarked savings account \( \hat{B} = B \) where the S&P 500 index is used as a benchmark, or a numéraire. The time period analyzed here covers data from 1 January 1980 to 31 December 2011.\(^5\) Note, the considered time period covers the most prominent crises discussed in the literature, which include the 1987 market crash, the LTCM and Russian crisis of 1998, the market decline in 2000 and the dot-com bubble of 2001, the recent subprime and financial crisis starting in June 2008.

In the following, we present the findings on parameter estimates and hence the

\(^5\)In the following, we discuss the results only for this time period. However, alternative time frames (with or without inclusion of crisis periods) have been used as a robustness check, and the results appeared to be quantitatively similar. The results are available from the authors upon request.
implications for the presence of money market bubbles discussed in Section 2. Furthermore, we discuss model selection using selection criteria presented in Section 4. Note however, model comparison is not the primarily goal of this study (which is to determine whether a money market bubbles is present or not), but merely serves as an indication for the best model under consideration given the information about the presence of money market bubbles; it could be used further for pricing of contingent claims.

Table 1 summarizes parameter estimates for different stochastic volatility model specifications. Note, the reported results are obtained using daily returns in percentages defined as $100 \times (\log(B_t) - \log(B_{t-1}))$. First of all, we find a large negative correlation $\rho$ between errors in the return and the variance equations, ranging from $-0.53$ to $-0.24$, indicating that the leverage effect is the strongest (weakest) for the Heston (3/2L) model specifications. The long-run mean of the variance $\theta$ ranges between 0.36 and 0.46 indicating long-run yearly volatility $\sqrt{252 \times \theta}$ to fall in a range between 9.5% and 10.8%. Note that $\kappa$ is not directly comparable between linear ($a = 0$) and nonlinear ($a = 1$) models, see [13]. Among linear model specifications $\kappa$ tends to be larger for Heston and Garch models, indicating greater speed of adjustment to the long-run variance level. For the nonlinear case (3/2N model) the speed of mean reversion is given by $\kappa V_t$. Finally, $\sigma$ is comparable for a given diffusion specification, i.e., for a given $b$, but not across different diffusion specifications, see [13]. For the 3/2N and 3/2L models one observes similar values for $\sigma$ corresponding to about 0.12. Note that our results are in line with those reported in the literature, see e.g. [20] and [13]$^6$, even if we

---

$^6$[13] use time frame from 1996 to 2004, which does not incorporate the most pronounced market crashes as in 1987 as well as the recent financial crisis. We have also used this time
work with the benchmarked savings account rather than with the S&P 500.\footnote{Remember, in order to make the results comparable, the paramater estimates have to be converted into the units used, as e.g., yearly values in [13].}

We now comment on the existence of money market bubbles. Note, under the Heston model the dynamics of $\hat{B}$ is always a martingale. We observe that for the $3/2N$, $3/2L$ as well as Garch models the inequalities formulated in Section 2 for presence of martingales, are violated, indicating that the price processes for these models resemble strict local martingales. As an implication of this results, we are now interested in which of the considered models provides the best fit to the data.

To decide on model performance, we first use visual inspection by looking at Figures 6.3 and 6.4, which show the QQ probability plot of residuals in the return and the variance equations, respectively. The QQ plot contrasts the quantiles of the estimated residuals with the quantiles from the standard normal distribution. A deviation of the residual data from the 45 degrees line indicates strong non-normality of residuals and thus, the evidence of misspecification. From the upper left panel of the figures, which correspond to the Heston model, we observe that the fat low tail for the return residuals (corresponding to negative returns) as well as the fat upper tail for the variance residuals (corresponding to high volatility) could not be captured appropriately. All alternative model specifications seem to perform similar (and outperform the Heston model) based on the QQ plot for return innovations (Figure 6.3), whereas the Garch model tends to slightly outperform all competing models in capturing outliers (high negative returns). The latter result becomes more pronounced in Figure 6.4 when examining variance period to estimate the parameters, and the results are consistent with [13].
residuals: here, the QQ plot for the Garch model shows no severe deviation from the 45 degrees line.

Secondly, we use the DIC statistic to rank the models. Comparing DIC values across different model specifications, we observe that the 3/2N model is ranked first among all models which allow for money market bubbles, whereas Heston model is ranked first overall. Note that this study considers stochastic volatility models without incorporating jump components. Although several studies treat jumps as secondary, if at all (see e.g. [4] and [20]), others find an inclusion of the jump component as essential (see e.g. [7], [20] and [11]). Naturally, including jumps (in returns or returns and volatility) might result in different ranking, and as documented e.g. in [30], the best performing stochastic volatility specification does not necessarily lead to the best performing overall specification when jump component is added to it. Since the objective of this study is to determine whether different model specifications allow strict local martingales (as supported by empirical evidence) and, we stay within the most parsimonious model specifications, that is, stochastic volatility (without jumps). However, our methodology can easily be extended to include jumps.

Finally, we compare model performance by their ability to capture observed volatility provided by VIX. Figure 6.2 shows the annualized daily spot volatility path $\sqrt{\nu_t}$ using the solid line for each model specification, whereas the dotted line represents the VIX volatility index. All volatility paths are shown on the same scale, going from 0% to 100% in annual terms. Although the overall pattern in volatility is similar across models, when volatility increases it tends to do so more

---

8 Note, for the applications we have in mind, that is, pricing under benchmark approach using the real world measure, we are interested in those models which allow strict local martingales.
sharply for the 3/2N and the 3/2L models, whereas the Heston model exhibits the least number of spikes in volatility compared to the remaining models.\(^9\) Thus, although all models appear to resemble VIX movements, the Heston model seems to underestimate spikes in VIX, which could be noticed even visually from the figure. This observation is confirmed when computing model errors reported in Table 2. Here, we present summary statistics for the relative error \(\delta_t\) defined as the percentage deviation from VIX, \(\delta_t = (\text{Vol}_t - \text{VIX}_t)/\text{VIX}_t\), its absolute value \(|\delta_t|\) as well as their sum of errors \(\sum_{t=1}^{T} \delta_t\) and \(\sum_{t=1}^{T} |\delta_t|\). We observe small negative values for the relative errors \(\delta_t\), ranging from to -0.46\% to -0.18\%, indicating that all models tend to slightly underestimate VIX. While Heston model under-performs all other models, the 3/2L and the Garch model perform equally well (based on the mean and the median statistics for \(\delta_t\)). This result also holds when \(\sum_{t=1}^{T} \delta_t\) is used as a decision criteria. Note, \(|\delta_t|\) ignores the direction of the deviation and does not cancel out under- and over-estimation effects. It leads to comparable values for all models allowing for money market bubbles (with 3/2L being slightly superior to the remaining models), whereas Heston model, which assumes \(\hat{B}\) to be a martingale, underperforms all models under consideration.

6 Conclusion

Bubbles have received a lot attention in the literature, though most of it has been focused on asset pricing bubbles, which assume the existence of an EMM. Money market bubbles exist if and only if an EMM fails to exist, which means that e.g. risk neutral pricing is not applicable. However, the benchmark approach, which

\(^{9}\text{These observations are in line with \cite{13} and \cite{30}.}\)
is formulated under the real world measure, is still applicable. Using stochastic volatility models well established in the literature, namely the Heston, $3/2$ with linear (L) and non-linear (NL) drift, and the Garch models, we test for the existence of money market bubbles in the US market. For this purpose, we use daily log-returns of the benchmarked savings account from 1 January 1980 to 31 December 2011. We find that all models which can produce money market bubbles in fact do so.

Comparing the performances of the different models, we observe that one out of three decision criteria (DIC) suggests that the Heston model fits data best, however, it does not support the empirical argument on the presence of money market bubbles. All alternative model specifications ($3/2$N, $3/2$L and Garch) outperform the Heston model, when using QQ-plots as well as model errors computed as relative deviations from the observed volatility (VIX) as the selection criteria. Comparing model performance across those which allow for money market bubbles, we tend to favor Garch as well as the $3/2$L model.

The results presented in this study lead to several ideas which could be interesting for further research. In particular, its implications can be used when pricing and hedging derivative contracts, put and call options, as well as long-dated insurance claims under the benchmark approach (i.e. directly under the real world measure).
### Table 1: Parameter estimates and their standard errors (in parentheses) for different stochastic volatility model specifications. The reported parameter estimates are obtained using data on the benchmarked (with S&P 500) savings account daily returns during time period from 1 January 1980 to 31 December 2011.

<table>
<thead>
<tr>
<th>Model</th>
<th>Heston (0.0;0.5)</th>
<th>3/2N (1.0;1.5)</th>
<th>3/2L (0.0;1.5)</th>
<th>Garch (0.0;1.0)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\kappa$</td>
<td>0.0609 (0.0053)</td>
<td>0.0347 (0.0037)</td>
<td>0.0117 (0.0040)</td>
<td>0.0531 (0.0059)</td>
</tr>
<tr>
<td>$\theta$</td>
<td>0.3604 (0.0237)</td>
<td>0.3696 (0.0426)</td>
<td>0.4660 (0.1100)</td>
<td>0.3860 (0.0202)</td>
</tr>
<tr>
<td>$\sigma_v$</td>
<td>0.1378 (0.0079)</td>
<td>0.1285 (0.0070)</td>
<td>0.1243 (0.0070)</td>
<td>0.1432 (0.0093)</td>
</tr>
<tr>
<td>$\rho$</td>
<td>-0.5356 (0.0325)</td>
<td>-0.4726 (0.0381)</td>
<td>-0.2462 (0.0349)</td>
<td>-0.4172 (0.0344)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>DIC</th>
<th>$p_D$</th>
<th>$\bar{D}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heston (0.0;0.5)</td>
<td>21436 (1)</td>
<td>2305</td>
<td>19131</td>
</tr>
<tr>
<td>3/2N (1.0;1.5)</td>
<td>21907 (2)</td>
<td>1776</td>
<td>20131</td>
</tr>
<tr>
<td>3/2L (0.0;1.5)</td>
<td>22946 (4)</td>
<td>694</td>
<td>22251</td>
</tr>
<tr>
<td>Garch (0.0;1.0)</td>
<td>22234 (3)</td>
<td>1495</td>
<td>20739</td>
</tr>
</tbody>
</table>

### Table 2: Model errors with ranking (in parentheses) for different stochastic volatility model specifications. The reported model are obtained using data on the benchmarked (with S&P 500) savings account daily returns during time period from 1 January 1980 to 31 December 2011.

<table>
<thead>
<tr>
<th>Model</th>
<th>Heston (0.0;0.5)</th>
<th>3/2N (1.0;1.5)</th>
<th>3/2L (0.0;1.5)</th>
<th>Garch (0.0;1.0)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary statistics for $\delta_t = (\text{Vol}_t - \text{VIX}_t)/\text{VIX}_t$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mean</td>
<td>-0.0046 (4)</td>
<td>-0.0043 (3)</td>
<td>-0.0018 (1)</td>
<td>-0.0020 (2)</td>
</tr>
<tr>
<td>std.dev</td>
<td>0.1800</td>
<td>0.1586</td>
<td>0.1572</td>
<td>0.1621</td>
</tr>
<tr>
<td>median</td>
<td>-0.0207 (3)</td>
<td>-0.0216 (4)</td>
<td>-0.0169 (2)</td>
<td>-0.0165 (1)</td>
</tr>
<tr>
<td>min</td>
<td>-0.5428</td>
<td>-0.6954</td>
<td>-0.5782</td>
<td>-0.4978</td>
</tr>
<tr>
<td>max</td>
<td>1.0486</td>
<td>1.1853</td>
<td>1.1367</td>
<td>0.9200</td>
</tr>
<tr>
<td>Summary statistics for $</td>
<td>\delta_t</td>
<td>=</td>
<td>\text{Vol}_t - \text{VIX}_t</td>
<td>/\text{VIX}_t$</td>
</tr>
<tr>
<td>mean</td>
<td>0.1403 (4)</td>
<td>0.1198 (2)</td>
<td>0.1186 (1)</td>
<td>0.1253 (3)</td>
</tr>
<tr>
<td>std.dev</td>
<td>0.1128</td>
<td>0.1040</td>
<td>0.1032</td>
<td>0.1028</td>
</tr>
<tr>
<td>median</td>
<td>0.1153 (4)</td>
<td>0.0986 (2)</td>
<td>0.0968 (1)</td>
<td>0.1035 (3)</td>
</tr>
<tr>
<td>min</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>max</td>
<td>1.0486</td>
<td>1.1853</td>
<td>1.1367</td>
<td>0.9200</td>
</tr>
</tbody>
</table>

**Sum of errors**

| | $\sum_{t=1}^{T} \delta_t$ | $\sum_{t=1}^{T} |\delta_t|$ |
|---|---|---|
| Heston (0.0;0.5) | -25.4774 (4) | 777.3140 (4) |
| 3/2N (1.0;1.5) | -23.5729 (3) | 663.5747 (2) |
| 3/2L (0.0;1.5) | -9.8642 (1) | 656.9741 (1) |
| Garch (0.0;1.0) | -11.2436 (2) | 694.2380 (3) |
Figure 6.2: Annualized daily spot volatility path $\sqrt{V_t}$ during time period from 1 January 1980 to 31 December 2011 for each stochastic volatility model specification: Heston (upper left), 3/2 (upper right), 3/2 with linear drift (lower left), Garch (lower right)
Figure 6.3: The QQ probability plot: plot quantiles from the return equation against quantiles from the normal distribution for each stochastic volatility model specification: Heston (upper left), 3/2 (upper right), 3/2 with linear drift (lower left), Garch (lower right). Data used in the estimation covers time period from 1 January 1980 to 31 December 2011.
Figure 6.4: The QQ probability plot: plot quantiles from the variance equation against quantiles from the normal distribution for each stochastic volatility model specification: Heston (upper left), 3/2 (upper right), 3/2 with linear drift (lower left), Garch (lower right). Data used in the estimation covers time period from 1 January 1980 to 31 December 2011.
References


7 Appendix: Description of the MCMC Algorithm

7.1 Model

The logarithm of the benchmarked savings account \( Y_t = \log \hat{B}_t \) solves

\[
\left( \frac{dY_t}{dV_t} \right) = \left( \begin{array}{c}
-\frac{1}{2} V_t \\
\kappa V_t (\theta - V_t)
\end{array} \right) dt + \left( \begin{array}{cc}
-V_t & 0 \\
0 & V_t^b
\end{array} \right) \left( \begin{array}{c}
1 \\
\rho \sigma \\
1 - \rho^2 \end{array} \right) \frac{dW_t}{\sqrt{\sigma}} \tag{7.18}
\]

where \( W(t) \) is a standard bivariate Brownian motion.

Time discretization of (7.18) (with \( \Delta > 0 \)) yields:

\[
Y_{tk+1} - Y_{tk} = R_{tk+1} = -\frac{1}{2} V_{tk} - \sqrt{V_{tk}} \varepsilon_{t+1}^y
\]

\[
V_{tk+1} - V_{tk} = \kappa V_{tk}^a (\theta - V_{tk}) + \sigma_v V_{tk}^b \varepsilon_{t+1}^v \tag{7.19}
\]

where \( \varepsilon_{t+1}^y = W_y(t+1) - W_y(t) \) and \( \varepsilon_{t+1}^v = W_v(t+1) - W_v(t) \) are standard normal random variables with correlation \( \rho \).

7.2 Priors

In order to simplify MCMC algorithm, we rely on conjugate priors for the model parameters. Those are chosen in line with the previous literature, see e.g. [20], [31] and [30] to guarantee a low level of information. In the following, we denote by \( N \) a normal distribution and by \( IG \) an inverse Gamma distribution.

- \( \theta \sim N(0, 1) \)
- \( \kappa \sim N(0, 1) \)
- \( \omega = \sigma_v^2 (1 - \rho^2); \omega \sim IG(2, 200) \)
- \( \psi = \sigma_v \rho; \psi | \omega^2 \sim N(0, 1/2\omega^2) \)
7.3 Posterior

The MCMC algorithm will consist of Gibbs and Metropolis-Hastings steps. Whenever possible, we use conjugate priors for the parameters in order to derive the complete conditionals. The only parameter (latent variable) which requires Metropolis-Hastings step is the variance. The posterior is given by the following expression

\[ p(\Theta, V | R) \propto p(R | \Theta, V)p(\Theta, V) = p(R | \Theta, V)p(V | \Theta)p(\Theta) \]

Note the following about the last expression

\[ p(R | \Theta, V)p(V | \Theta) \propto \prod_{t=1}^{T} p(R_t | V_t, V_{t-1}, \Theta)p(V_t | V_{t-1}, \Theta) = \prod_{t=1}^{T} p(R_t, V_t | V_{t-1}, \Theta) \]

since \( p(R | \Theta, V) = \prod_{t=1}^{T} p(R_t | V_t, V_{t-1}, \Theta) \) by conditional independence and \( p(V | \Theta) \propto \prod_{t=1}^{T} p(V_t | V_{t-1}, \Theta) \) by Markov property.

From equation (7.19) we see that the joint distribution is bivariate normal with the following parameters

\[
\begin{align*}
\mu &= \left(\begin{array}{c}
-\frac{1}{2}V_{t-1} \\
V_{t-1} + \kappa V_{t-1}^a (\theta - V_{t-1})
\end{array}\right) \\
\Sigma &= \left(\begin{array}{cc}
V_{t-1}^{-1} & -\rho \sigma_v V_{t-1}^{b+0.5} \\
-\rho \sigma_v V_{t-1}^{b+0.5} & \sigma_v^2 V_{t-1}^{2b}
\end{array}\right) \\
&= \left(\begin{array}{cc}
V_{t-1}^{-1} & -\rho \sigma_v V_{t-1}^{b+0.5} \\
-\rho \sigma_v V_{t-1}^{b+0.5} & \sigma_v^2 V_{t-1}^{2b}
\end{array}\right)
\end{align*}
\]

(7.20)

Given normality of the joint distribution we can determine the conditional distribution \( p(R_t | V_t, V_{t-1}, \Theta) \). It is also normal with parameters given by

\[
\begin{align*}
\mu_{R_t|V_t} &= -\frac{1}{2}V_{t-1} - \rho \sigma_v^{-1}V_{t-1}^{0.5-b}(V_t - V_{t-1} - \kappa V_{t-1}^a (\theta - V_{t-1})) \\
\sigma_{R_t|V_t}^2 &= V_{t-1}(1 - \rho^2)
\end{align*}
\]

(7.21)

On the other hand, for \( p(V_t | R_t, V_{t-1}, \Theta) \) we obtain a normal distribution with

\[
\begin{align*}
\mu_{V_t|R_t} &= V_{t-1} + \kappa(V_{t-1})^a(\theta - V_{t-1}) - \rho \sigma_v V_{t-1}^{b-0.5}(R_t + \frac{1}{2}V_{t-1}) \\
\sigma_{V_t|R_t}^2 &= \sigma_v^2 V_{t-1}^{a+b}(1 - \rho^2)
\end{align*}
\]
We can rewrite the posterior as

\[ p(\Theta, V | R) \propto \prod_{t=1}^{T} p(R_t, V_t | V_{t-1}, \Theta)p(\Theta) \]

Since we cannot sample from the posterior we have to construct the complete conditionals, in order to use MCMC. The algorithms to draw from the complete conditional in each iteration is outlined below. Most of the complete conditional distributions can be derived using standard results in, e.g., [10].

- **Drawing \( \theta \):** The complete conditional distribution has the form of a regression with

  \[ y_t = \left[ V_t - V_{t-1} + \kappa V_{t-1}^{a+1} + \rho \sigma_v V_{t-1}^{b-0.5}(R_t + \frac{1}{2} V_{t-1}) \right] / \left[ \sigma_v V_{t-1}^{b} \sqrt{1 - \rho^2} \right] \]

  as a dependent variable and

  \[ x_t = [\kappa V_{t-1}^a] / \left[ \sigma_v V_{t-1}^{b} \sqrt{1 - \rho^2} \right] \]

  as an explanatory variable. The regression has a known error variance of one.

- **Drawing \( \kappa \):** The complete conditional distribution has the form of a regression with

  \[ y_t = \left[ V_t - V_{t-1} + \rho \sigma_v V_{t-1}^{b-0.5}(R_t + \frac{1}{2} V_{t-1}) \right] / \left[ \sigma_v V_{t-1}^{b} \sqrt{1 - \rho^2} \right] \]

  as a dependent variable and

  \[ x_t = [\kappa V_{t-1}^a \theta - V_{t-1}] / \left[ \sigma_v V_{t-1}^{b} \sqrt{1 - \rho^2} \right] \]

  as an explanatory variable. The regression has a known error variance of one.

- **Drawing \( \rho \) and \( \sigma_v \):** We follow [31] and use reparameterization \( \psi = \sigma_v \rho \) and \( \omega^2 = \sigma_v^2 (1 - \rho^2) \). This yields to the regression with

  \[ y_t = \left[ V_t - V_{t-1} - \kappa V_{t-1}^a \theta - V_{t-1} \right] / \left[ V_{t-1}^b \right] \]

  and

  \[ x_t = -V_t^{-1/2}(R_t + \frac{1}{2} V_{t-1}). \]

  The unknown homoscedastic error variance is given by \( \omega^2 = \sigma_v^2 (1 - \rho^2) \). Using this setup, we are able to estimate \( \psi \) and \( \omega^2 \). The parameters of interest are then given by \( \sigma_v^2 = \omega^2 + \psi^2 \) and \( \rho = \psi / \sigma_v \).
**Drawing** $V_t$: Variance $V_t$ is a latent variable which does not have recognizable complete conditional and therefore, we rely on the Metropolis-Hastings algorithm. As outlines in Section 3, the complete conditional distribution for $V_t$ can be written as

$$p(V_t|\Theta, R) \propto p(R_{t+1}, V_{t+1}|V_t, \Theta)p(V_t|R_t, V_{t-1}, \Theta),$$

where $p(R_{t+1}, V_{t+1}|V_t, \Theta)$ follows bivariate normal distribution and $p(V_t|R_t, V_{t-1}, \Theta)$ follows univariate normal distribution. The Metropolis-Hasting step proposes a new variance $V_t^{(i)}$ in $i^{th}$ draw by drawing from $p(V_t|R_t, V_{t-1}, \Theta)$ and accepting that draw with probability

$$\min \left\{ \frac{p(R_{t+1}, V_{t+1}|V_t^{(i)}, \Theta)}{p(R_{t+1}, V_{t+1}|V_t^{(i-1)}, \Theta)}, 1 \right\}.$$