Regime heteroskedasticity in Bitcoin: A comparison of Markov switching models

Chappell, Daniel

Birkbeck College, University of London

28 September 2018

Regime heteroskedasticity in Bitcoin: A comparison of Markov switching models

Daniel R. Chappell
Department of Economics, Mathematics and Statistics
Birkbeck College, University of London
research.chappell@gmail.com

28th September 2018

Abstract

Markov regime-switching (MRS) models, also known as hidden Markov models (HMM), are used extensively to account for regime heteroskedasticity within the returns of financial assets. However, we believe this paper to be one of the first to apply such methodology to the time series of cryptocurrencies. In light of Molnár and Thies (2018) demonstrating that the price data of Bitcoin contained seven distinct volatility regimes, we will fit a sample of Bitcoin returns with six m-state MRS estimations, with $m \in \{2, ..., 7\}$. Our aim is to identify the optimal number of states for modelling the regime heteroskedasticity in the price data of Bitcoin. Goodness-of-fit will be judged using three information criteria, namely: Bayesian (BIC); Hannan-Quinn (HQ); and Akaike (AIC). We determined that the restricted 5-state model generated the optimal estimation for the sample. In addition, we found evidence of volatility clustering, volatility jumps and asymmetric volatility transitions whilst also inferring the persistence of shocks in the price data of Bitcoin.

Keywords

Bitcoin; Markov regime-switching; regime heteroskedasticity; volatility transitions.

List of Tables

List of Figures

Contents

1 Introduction

Cryptocurrencies are digital "assets"[1] motivated to act as a medium of exchange, supposedly outside the influence of any government or central bank. They are based upon high-level cryptography that is applied to not only secure and verify transactions, but to also modulate the creation of additional coins from a finite issue. As more coins are "mined", the remaining coins become more costly to decrypt, thereby providing an artificial scarcity and driving a perceived value based upon rarity and cost. The genesis of the cryptocurrency gold rush can be traced back to the seminal paper published under the pseudonym Satoshi Nakamoto in 2008. Various individuals have either been accused of being the elusive Nakamoto, or have stood up to pronounce "I'm Nakamoto". However, mystery still surrounds the true identity of the individual, or group, who first proposed a peer-to-peer (P2P) transactional system based upon the Blockchain distributive-ledger technology. The first cryptocurrency born out of Nakamoto's paper was Bitcoin in 2009. Since then thousands of other cryptocurrencies have been launched, with the current total capitalisation of the cryptocurrency market at over US$202.5 billion[2].

These pseudo-currencies have not been far from the headlines over the past year due to the exceptional bubble, and subsequent crash, in the price of Bitcoin towards the end of 2017. Since then, the value of Bitcoin, and many other cryptocurrencies, have drifted slowly lower through most of 2018. However, as Mark Twain expressed "History does not repeat itself, but it often rhymes." We have seen this all before, in the bubble and crash of Bitcoin in 2013. Retail investors will lick their wounds as a result of the most recent crash but they will probably not learn their lessons. In addition, crytocurrencies have also made headlines with respect to crypto exchanges being hacked, assets being stolen[3] and in some cases the exchange itself collapsing into bankruptcy[4]. The systematic and systemic risk surrounding these unregulated instruments cannot be ignored. The U.S. Securities and Exchange Commission (SEC) has warned of a possible inability to "pursue bad actors or recover funds"[5] due mostly to the decentralised and unregulated nature of the cryptocurrency marketplace.

[1] A number of papers have been written on whether cryptocurrencies are a method of exchange (currency) or an investment vehicle (asset), see Kubát (2015).

[2] Total capitalisation of the cryptocurrency market stands at US$202,513,161,701. Available at: www.coinmarketcap.com. Last accessed: 12:00 UTC, 16th September 2018.

[3] In 2018 alone, the exchange Bithumb was hacked for approximately £31 million, whilst the Coincheck exchange suffered a hack in the region of £408 million. Available at: www.coindesk.com/bithumb-exchanges-31-million-hack-know-dont-know/. Last accessed: 9th September 2018.

[4] The exchange Youbit was hacked for a second time in 2017 with the theft of 17% of total assets, resulting in the exchange declaring bankruptcy. Available at: www.coindesk.com/south-korean-bitcoin-exchange-declare-bankruptcy-hack/. Last accessed: 9th September 2018.

[5] Available at: www.sec.gov/news/public-statement/statement-clayton-2017-12-11. Last accessed: 16th September 2018.

In addition, the UK's Financial Conduct Authority (FCA) issued advice on the risks of investing in cryptocurrencies[6], citing: leverage; charges; funding costs; price transparency; and price volatility as the the major risks associated with trading such instruments. Whilst the authors of this paper do not promote investment in the current unregulated crypto market, we do expect that regulated cryptocurrencies will appear in the future, supported by the benefits of Blockchain for verifying the provenance of digital-assets.

Regime change in financial markets

Campbell et al. (1997) argued that for financial time series, the assumption of constant volatility over some period of time was statistically inefficient and logically inconsistent. This was due to the fact that such time series tended to demonstrate volatility clustering, i.e. large (small) returns were typically followed by large (small) returns. This was one of three stylised facts often quoted when discussing financial time series, that were first stated by Mandelbrot (1963a,b and 1967). The other two facts presented were the heteroskedasticity of variance and the non-normal leptokurtic distribution of financial asset returns. An additional stylised characteristic exhibited by financial time series is the "leverage effect". This term evolved over time but its etymological roots can be found in Black (1976) and refers to the asymmetric phenomena of negative returns presenting higher volatility than positive returns of the same magnitude. Reasons proposed for the presence of volatility clustering and the heteroskedasticity of variance in financial time series include: dependence on the rate of information arrival to the market (Lamoureux and Lastrapes, 1990); errors in the learning processes of economic agents (Mizrach, 1990); and the artificial nature of a calendar timescale in lieu of a perceived operational timescale (Stock, 1990).

The manifestation of these stylised facts in financial markets, was that such markets tended to change their behaviour in a sudden manner. The duration for which the new behaviour persists was generally unknown *ab intra*. Ang and Timmermann (2012) stated that regime switching models were able to capture sudden changes in the price dynamics of financial assets, where such changes arose as a result of the aforementioned inherent stylised characteristics. They continued; whilst regime switching models could be used to identify past categorical delineations in time series data, where regimes were found econometrically, they could also be used for *ex-ante* forecasting and optimising portfolio choice in real-time. In addition, Ang and Timmermann stated that regime switching models were able to accommodate jumps in financial time series, by simply considering each jump as a special regime that is exited in the immediately following instance. Given the published work of Molnár and Thies (2018) into the price volatility of Bitcoin, we expect to see evidence of such volatility jumps in the results of our MRS estimations.

[6]Available at: www.fca.org.uk/news/news-stories/consumer-warning-about-risks-investing-cryptocurrency-cfds. Last accessed: 16th September 2018.

The first application of a regime switching model can be found in Hamilton (1989). The model related to the business cycle moving from states of expansion to recession and back again around a long-term trend. Since then, regime change models have been applied to a variety of financial applications including equities by Pagan and Sosssounov (2003). Meanwhile, Sims and Zha (2006) fitted U.S. data to a multivariate regime-switching model for monetary policy. They used a four-state regime-switching model with time-varying coefficients to capture changes in policy rule. They found their model's fit to be superior to all other models compared that also permitted dynamic coefficients.

Regime switching models have also been applied to non-financial scenarios. Smith, Sola and Spagnolo (2000) used a discrete state regime-switching model to estimate the transition probabilities of a simple four-state model. The model aimed to emulate the arms race between Turkey and Greece over the period 1958–1997. The four states represented whether each country chose a high or low share of military expenditure, with pay-offs assumed to match those of the Prisoner's dilemma. They found that translating a relative simple two-by-two game theory exercise into an empirical model was a complicated process, requiring 19 free parameters. Although the number could be reduced by making basic assumptions with regard to the strategy of each player. Similarly, we will examine the transition probabilities within the unrestricted MRS estimations used in this paper and make the basic assumption that near-zero values are indeed zero values. To do so we will fit transition restriction matrices when required. This will decrease the complexity of the estimation process and purportedly increase the goodness-of-fit of the results.

Chang et al. (2017) proposed a new approach to modelling regime switching. Their model incorporated an autoregressive latent factor that determined regimes dependent on whether some threshold level was breached. As opposed to previous models, they permitted the latent factor to be endogenous with the innovations of the observed sequence rather than exogenous; which would otherwise have transformed their approach back into the conventional Markov regime-switching model.

Markov regime-switching models

Markov regime-switching (MRS) models, or hidden Markov (HMM) models, assume that an observed process is motivated by an unobserved state process. Such models are a special form of dependent mixture model, consisting of two parts: a parameter process that satisfies the Markov property; and a state-dependent process, that results in the distribution of specific observations being dependent only on the current state and not on previous observations or states (Langrock et al., 2016). The mathematical grounding of these models was first developed by Baum and Petrie (1966). They used a Markov process to simulate the hidden sequence by which an observed sequence was generated; whilst using maximum likelihood (ML) methodology to estimate the unknown parameters of the transition and

observation probability matrices. The computation of the likelihood, L_T, of T sequential observations $(x_1, x_2, x_3, ..., x_T)$ for an m-state Markov regime switching model should require Tm^T operations. However, the derivation of a convenient formula for the likelihood of such models requiring only Tm^2 operations can be found in Langrock et al. (2016). They additional stated that, HMMs are perfectly suited to handling data that is overdispersed and serial dependent.

Four years after Baum and Petrie first published on HMMs, Baum et al. (1970) published a paper demonstrating the solution for a single observation sequence. They recommended the application of such models for capturing stock market behaviour and weather forecasting. However, it was a further 13 years before a solution for multiple observation sequences was published by Levinson et al. (1983). To do so, they had to develop the "left-to-right HMM" and also assume independence between each observation sequence.

Almost two decades later, the restrictive assumption of independence in the multiple observation sequences' framework was dropped by Li et al. (2000). In doing so, they also identified two types of multiple observation sequences, namely: the uniform dependence observation sequences and the independence observation sequences. A year later, Ghahramani (2001) set out an extensive tutorial linking HMMs and Bayesian Networks, thereby enabling new generalisations of MRS models using: multiple unobservant state variables; and combined continuous and discrete variables.

Over the last decade, MRS models have been applied to the time series of more traditional financial assets, including: forecasting stock prices (Hassan and Nath, 2005); applications in foreign exchange (Idvall and Jonsson, 2008); forecasting S&P daily prices within a equity-selection strategy (Lajos, 2011); predicting regimes in inflation indexes (Kritzman, Page and Turkington, 2012); analysing trends in equity markets (Kavitha, Udhayakumar and Nagarajan, 2013); and selecting stocks based on predicting future regimes (Nguyen and Nguyen, 2015). However, we believe this paper to be one of the first to use MRS methodology to account for regime heteroskedasticity in the price volatility of cryptocurrencies.

A fool's errand

Since 2009, interest in cryptocurrencies has experienced exponential growth, buoyed by the changing demographics of society with respect to the acceptance of disruptive innovation in financial technology (FinTech). As such, efforts into modelling the conditional moments of the flagship cryptocurrency Bitcoin have been extensive. There have been numerous studies published into finding the optimal single-regime generalised autoregressive conditional heteroskedasticity (GARCH) model for Bitcoin. These include: Glaser et al. (2014) and Gronwald (2014), who declared the linear GARCH model was superior; Bouoiyour and Selmi (2015) and Dyhrberg (2016a), who claimed that it was the Threshold GARCH

(TGARCH) variant that was optimal; whilst Katsiampa (2017) found that both the long-run and short-run memory elements of the Component GARCH (CGARCH) variant made it optimal for modelling the conditional variance of Bitcoin.

The simplifying assumption that a single-regime GARCH model is suitable for capturing the price risk of a cryptocurrency is simply ill-founded. Regime heteroskedasticity has been shown to be present in the time series of many financial asset returns, including Bitcoin by Molnár and Thies (2018). They utilised a Bayesian change point model to detect structural changes in the cryptocurrency and then categorised partitions of the entire time series into one of seven volatility regimes. Bauwens et al. (2010, 2014) shown that if single-regime GARCH models were applied to time series that contained structural breaks, then the estimates tended to be biased and the forecasts inferior. Therefore, searching for the optimal single-regime GARCH variant for Bitcoin may well have been a fool's errand for the authors mentioned previously.

The pressure to publish has led many authors to present poorly conceived papers into the optimal single-regime GARCH variant for estimating and forecasting the conditional variance of cryptocurrencies. For examples, see Katsiampa (2017) or Chu et al. (2017). Notwithstanding her apparent confusion between daily growth rates and daily log returns, Katsiampa identified the Component GARCH (CGARCH) variant as optimal for the estimation of the conditional variance of Bitcoin. This finding was under the ill-conceived assumption of conditionally normally distributed innovations. However, given the fundamental definition of the GARCH model, Sun and Zhou (2014) stated that the distribution of innovations fits hand-in-glove with the conditional distribution of future returns. In addition, Bai et al (2003) stated that the resultant kurtosis associated with the assumption of Gaussian innovations, tended to significantly underestimate the observed kurtosis.

A more straightforward approach would be to assume Student's t-distributed innovations, since the distribution possesses thicker tails than the Gaussian distribution, especially when the degrees of freedom are low. However as Shaw (2018) stated, the Student's t-distribution does not possess a moment generating function (MGF). Therefore, applying Student's t-distributed innovations within a risk-neutral framework for financial engineering purposes, could result in a call option that possessed infinite value (Shaw, 2018). Similarly, Shaw argued that the lack of an MGF for the skewed generalised error distribution (SGED) in certain circumstance meant that this was also a poor solution to the issue of identifying an accurate and robust fix for the assumption of the distribution of innovations. Concluding, Shaw demonstrated that the innovations of a simple single-regime linear GARCH(p, q) model, with $p, q \in \{1, ..., 5\}$, for six major cryptocurrencies including Bitcoin, were indeed conditionally non-normally distributed. This was achieved by applying a Kolmogorov-type non-parametric test to eliminate the possibility of Gaussian innovations.

Markov regime-switching GARCH models

Hamilton and Susmel (1994), as motivation for developing their hybrid MRS and ARCH model (SWARCH), stated that the ARCH model provided relatively poor forecasts. Their aim was to address the spurious high persistence issues that arose from using ARCH models on samples that contained distinct volatility regimes. They argued that being able to model regime switching along with conditional heteroskedasticity, would allow for the capture of changes in the factors that affect volatility and overall reflect the changing nature of market conditions. Over the past two decades, MRS-GARCH models have been used to estimate the conditional variance of more traditional financial returns time series, including: stock index returns in Marcucci (2005); commodity returns in Alizadeh et al. (2008); stock returns in Henry (2009); and exchange-rate returns in Wilfling (2009).

In recent days, Ardia et al. (2018) released a research note that combined the MRS and GARCH methodologies under the framework of conditionally non-normally distributed innovations for Bitcoin [Regime changes in Bitcoin GARCH volatility dynamics. *Finance Research Letters*. Available online: 10th August 2018]. Whilst they were able to identify the 2-regime MRS-GARCH model as superior to the single-regime version within the comparative frame of forecasting value-at-risk (VaR), they limited their research to the 2- and 3-state MRS-GARCH models. A major issue with 2-state MRS models is that they do not allow for volatility jumps; transitions can only occur between the two adjoining states and not to additional disjoint states as is the case with higher dimensional MRS models.

As such, we recoil from the myopic pursuit of identifying the optimal single-regime GARCH variant for Bitcoin for the reasons already stated above, see *A fool's errand*. Instead, and in response to Ardia et al. (2018) and Molnár and Thies' (2018) papers, **we fit six m-state Markov regime switching (MRS) models, with $m \in \{2, ..., 7\}$, to the price data of Bitcoin, in order to identify the optimal number of states for modelling the cyrptocurrency's regime heteroskedasticity**. We will use maximum likelihood (ML) for fitting the models and goodness-of-fit between the estimations will be judged using three negated information criteria, namely: Bayesian (BIC); Hannan-Quinn (HQ); and Akaike (AIC).

Following on: Section 2 introduces the data; Section 3 presents the models; Section 4 states the results; and Section 5 concludes.

2 Data

For our analysis of the cryptocurrency Bitcoin, we elected to use the daily closing prices of the Bitcoin Coindesk Index for our data. The data is available to the public at www.coindesk.com/price. The sample accessed was from 22nd April 2014 to 31st May 2018. As such, the data from the 1st June 2018 to 28th September 2018 can be used for future research as an out-of-sample data set for the evaluation of the forecasting properties of the optimal model within both VaR and ES frameworks. However, an examination of the runtimes for the higher-state estimations identified, see Section *4.7 Estimation runtimes*, meant that an academically-rigorous forecasting exercise for the seven MRS models was beyond both the programming skills of the authors and the scale of this paper.

Since Bitcoin is traded seven days a week, the data set contained 1,501 observations. The data was plotted to check for outliers and the date stamp of each observation was examined for any repetition within the set. To limit the impact of outliers in our data, we examined the daily returns of the sample within a logarithmic framework. Using the downloaded closing price data, as measured at 00:00 UTC each day, daily log returns were found by taking the natural logarithm of the ratio of two consecutive closing prices. The new sample set of log returns contained 1,500 observations. Table 1 presents the summary statistics for the sample of daily log returns for Bitcoin. A kurtosis value of 8.284923 indicated the presence of non-normal leptokurtic behaviour. Figures 1, 2 and 3 (overleaf) illustrate respectively: the daily closing price; the daily log returns; and the frequency distribution of daily log returns for the Bitcoin Coindesk Index sample.

Table 1. Summary statistics for Bitcoin (23rd April 2014 to 31st May 2018)

Statistic	Daily log returns
Observations	1,500
Mean	0.001825
Median	0.001820
Maximum	0.226412
Minimum	-0.247132
Std.Dev.	0.038714
Skewness	-0.281471
Kurtosis	8.284923
VaR 1% (loss)	11.6309%
VaR 5% (loss)	6.4756%

Figure 1. Daily closing price, Bitcoin Coindesk Index (22nd April 2014 to 31st May 2018)

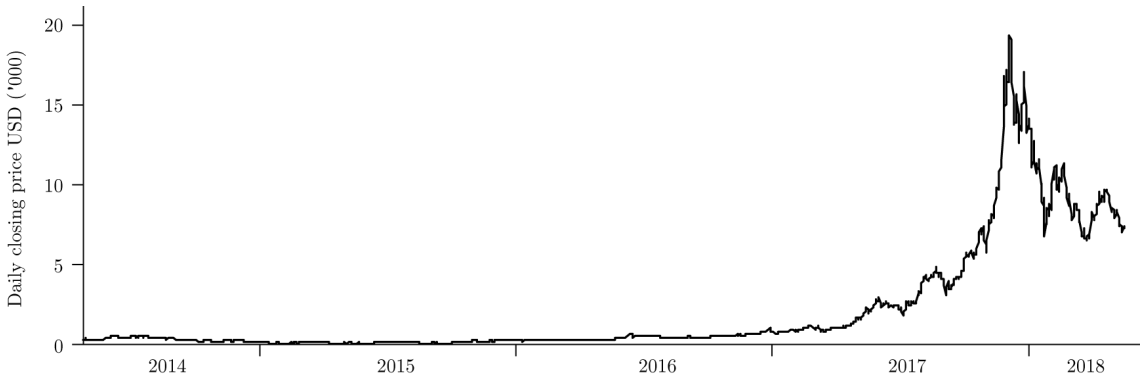


Figure 2. Daily log returns, Bitcoin Coindesk Index (23rd April 2014 to 31st May 2018)

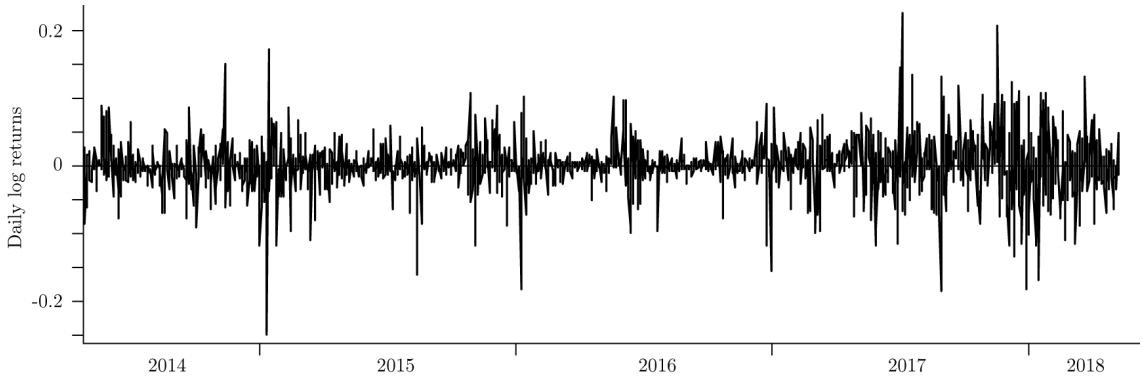
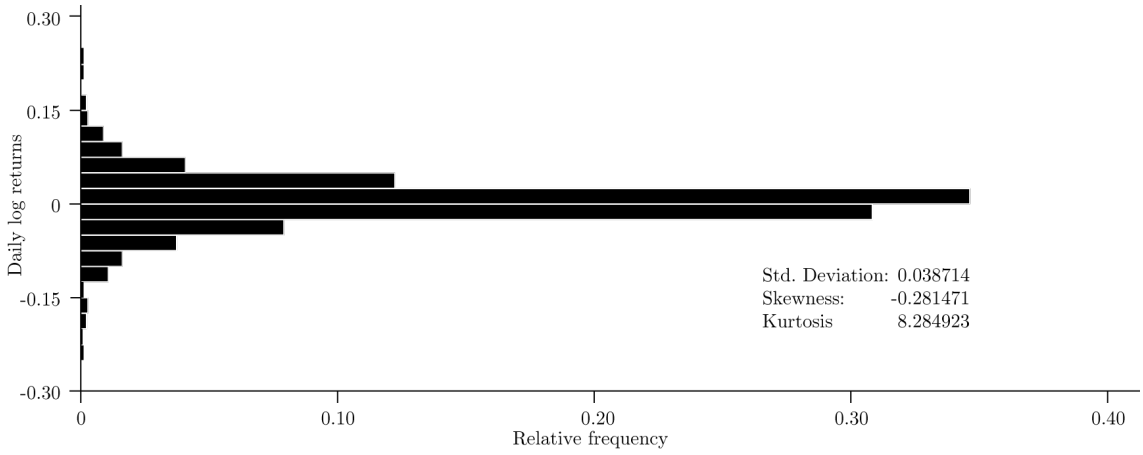


Figure 3. Frequency distribution of daily log returns (23rd April 2014 to 31st May 2018)



Jarque-Bera (JB) and the more powerful Kolmogorov-Smirnov (KS) tests were applied and the null hypotheses of a normally distributed sample were strongly rejected at the 1% significance level for both tests. The unit root tests, Augmented Dickey-Fuller (ADF) and Phillips-Peron (PP), were conducted and the null hypotheses of a unit root in the returns were rejected in both tests at the 1% significance level, indicating that stationarity was present. The results of the tests on the sample are presented in Table 2.

Table 2. JB, KS, ADF and PP test results for Bitcoin (23rd April 2014 to 31st May 2018)

Test	Score (significance)
Jarque-Bera	1,765.457***
KS	0.44861***
ADF	-38.22321***
PP	-38.22714***

* significance at 10% level, ** significance at 5% level and *** significance at 1% level.

Table 3 states the number of positive and negative observations within the sample as well as the mean of the positive and negative observations. The values indicated that there were 124 fewer negative observations in our sample. This was expected given Bitcoin increased in value by $1,345.57\%$ over the interval sampled (US\$484.43 on 22nd April 2014 to UD\$$7,487.19$ on 31st May 2018). However, the overall skewness of the sample was negative, see Table 1. This suggested that the magnitude of the fewer daily losses must have exceeded those of the more frequent daily gains. A fact that is reinforced by comparing the mean of the positive observations with the mean of the negative observations in Table 3. These facts suggests the presence of the asymmetric volatility phenomenon (AVP) or leverage effect, in the price data of Bitcoin.

Table 3. Summary statistics for the positive and negative return subsamples

Test	Score (significance)
Total obs.	1,500
Positive obs.	812
Mean positive obs.	0.025149
Negative obs.	688
Mean negative obs.	-0.025702

The following section introduces the methodology applied in this paper.

3 Methodology

Markov regime-switching (MRS) models, or hidden Markov (HMM) models, assume that an observed process is motivated by an unobserved state process. Such models are a special form of dependent mixture model, consisting of two parts: a parameter process that satisfies the Markov property; and a state-dependent process, that results in the distribution of specific observations only depending on the current state and not on previous observations or states (Langrock et al., 2016). For simplicity we assume the error terms in our regression models are *iid* normally distributed and conditional on the current regime and that the Markov model possesses constant probabilities, such that H_{t-1} is a constant.

3.0.1 Regime probabilities

The first-order Markov assumption necessitates the probability of being in a specific state depends on the the previous state:

$$P(s_t = j \mid s_{t-1} = i) = p_{ij}(t)$$

Where $p_{ij}(t)$ is time-invariant, so:

$$p_{ij}(t) = p_{ij} \quad , \forall t$$

As such, for an m-state model, we can construct an m x m transition matrix, $p(t)$, where the entry $p_{ij(t)}$ corresponds to the transition probability from state i to state j at time t:

$$p(t) = \begin{bmatrix} p_{11}(t) & p_{12}(t) & \cdots & p_{1m}(t) \\ p_{21}(t) & p_{22}(t) & \cdots & p_{2m}(t) \\ \vdots & \vdots & \ddots & \vdots \\ p_{m1}(t) & p_{m2}(t) & \cdots & p_{mm}(t) \end{bmatrix}$$

Within the transition probability matrix, each row specifies a complete set of conditional probabilities, as such a separate multinomial specification is generated for each row:

$$p_{ij}(H_{t-1}, \gamma_i) = \frac{e^{(H'_{t-1}\gamma_{ij})}}{\sum_{\gamma=1}^{m} e^{(H'_{t-1}i\gamma)}}$$

Where $i \in \{1, ..., m\}$ and $j \in \{1, ..., m\}$.

3.0.2 Likelihood evaluation and filtering

The necessity for a Markov switching model to have the state at time t be dependent on the previous state at time $t-1$, creates a serious issue in the estimation of the parameters for the mixture model. Consider the pdf of a basic 2-state mixture model:

$$\mathbb{P}(S_t = 0) \cdot \text{pdf}(X_t \mid (S_t = 0)) + \mathbb{P}(S_t = 1) \cdot \text{pdf}(X_t \mid (S_t = 1))$$

This can be rewritten as:

$$p \cdot \text{pdf}(X_t \mid (S_t = 0)) + (1 - p) \cdot \text{pdf}(X_t \mid (S_t = 1))$$

The probability of observing X_t is not dependent on any other observation of X, i.e. X_{t-1}. However, the Markov switching model requires knowledge of the previous unobserved interval's state. This necessity for Markov switching models means that the above expressions are insufficient for the maximum-likelihood estimation. Hamilton (1994) solved this issue by making the following substitutions:

$$\mathbb{P}(S_t = 0) = \mathbb{P}(S_{t-1} = 0 \mid \{X_{t-1}, ..., X_1\}) \cdot p + \mathbb{P}(S_{t-1} = 1 \mid \{X_{t-1}, ..., X_1\}) \cdot (1 - q)$$

$$\mathbb{P}(S_t = 1) = \mathbb{P}(S_{t-1} = 0 \mid \{X_{t-1}, ..., X_1\}) \cdot (1 - p) + \mathbb{P}(S_{t-1} = 1 \mid \{X_{t-1}, ..., X_1\}) \cdot q$$

Hamilton then stated that the above expressions can be solved by calculating:

$$\mathbb{P}(S_{t-1} = 0 \mid \{X_{t-1}, ..., X_1\})$$

from:

$$\mathbb{P}(S_{t-2} = 0 \mid \{X_{t-2}, ..., X_1\})$$

As such, Hamilton demonstrated that the likelihood function of X_t can be recursively calculated from X_{t-1}. However, the process must be specified at time 0. That is, the initial regime probabilities must be specified.

3.0.3 Initial regime probabilities

The Markov regime-switching filter requires populating the filtered regime probabilities at time 0. This can be achieved in one of four ways:

- supplying known regime probability values (user specified);
- assuming equally likely regimes at the outset (uniform);
- estimate the probabilities as if they were parameters themselves (estimated); or
- assume the values are functions of the parameters that generate the transition matrix (ergodic solution).

For simplicity we used the last option for the initial regime probabilities due to the other three options being either unlikely (uniform) or not possible (user supplied). However, the ergodic solution can lead to arbitrary initialisations for time-varying transition probabilities, which was not an issue for this paper.

3.0.4 Probability transition smoothing

In estimating the probability transitions, we opted to apply a smoothing technique that uses all of the information in the sample, $\mathbb{I}_T$, and not just the information that preceded period t and the observation at time t itself, $\mathbb{I}_t$:

$$\mathbb{P}(S_t = 0 \mid \{X_T, ..., X_1\})$$

instead of just:

$$\mathbb{P}(S_t = 0 \mid \{X_t, ..., X_1\})$$

Hamilton (1989) referred to the function used to produced smoothed probability transitions as a "full-sample smoother". However, this function requires $(T^2/2) + 1$ computations per set of observations from time 1 to time T. Five years later, Kim (1994) set out a simpler method than the one used by Hamilton that reduced the number of computations required from T^2 to $2T$. This revision was based upon the fact that smoothed and *ex-post* probabilities are equal for the last observation, as such a recursive formula could be applied:

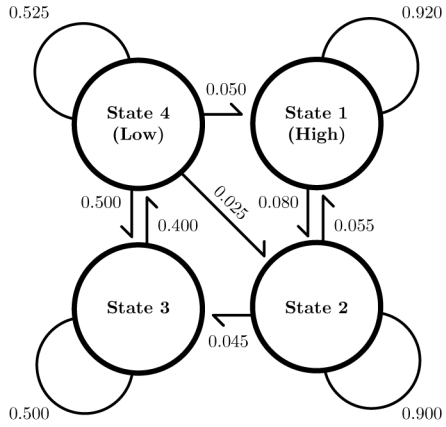
$$\mathbb{P}(S_t = 0 \mid X_T) = \mathbb{P}(S_t = 0 \mid X_t) \cdot \left(p_{00} \frac{\mathbb{P}(S_{t+1} = 0 \mid X_T)}{\mathbb{P}(S_{t+1} = 0 \mid X_t + 1)} + p_{10} \frac{\mathbb{P}(S_{t+1} = 1 \mid X_T)}{\mathbb{P}(S_{t+1} = 1 \mid X_t + 1)} \right)$$

The improved estimates for the smoothed regime transitions are a result of the Markov transition probabilities connecting the likelihood of observations from different instances.

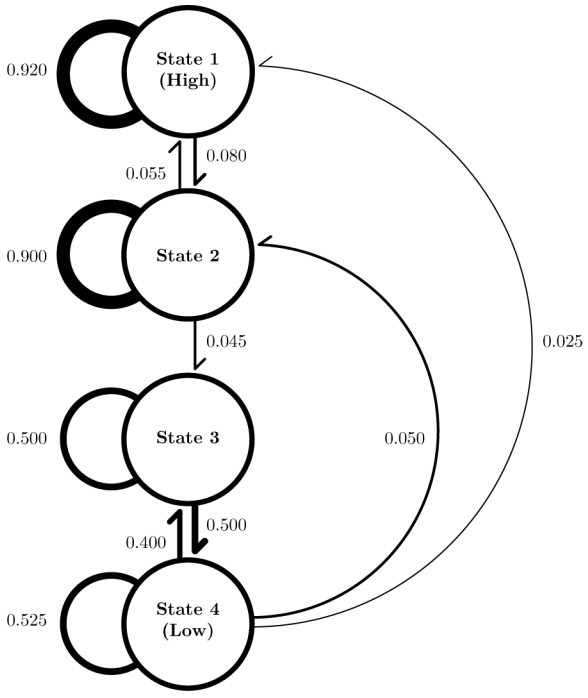
3.0.5 Transition probability diagrams

Traditionally, transition probability diagrams have been looped or longitudinal in nature to express the interconnectivity of different states within a process. However, the states referenced in this paper not only possess a categorical delineation but also an ordered delineation; low-to-high volatility regimes via a variable number of transitions. As such, we developed a column transition probability diagram in order to more intuitively represent the ordinal nature of the volatility regimes examined. As a result, volatility jumps, which involve a transition away from one state to another that is neither directly above nor below itself can be identified more easily in the proposed vertically-orientated transition probability diagram.

In addition, we communicate the magnitude of a probability transition by not only stating the value next to the corresponding transition, but also adjusting the weight of the stroke used for each individual transition in proportion to the magnitude of the transition probability. An example of both diagram styles mapping the same fictitious process can be seen below. It is more clearly illustrated in the vertically-orientation representation, that the process permits jump transitions from State 4 to both States 1 and 2.



Traditional loop orientation



Proposed vertical orientation

3.0.6 Transition restriction matrices

Once an initial unrestricted m-state MRS estimation was complete, we investigated the resultant transition probability matrix. If we found transition probability values that were near-zero in the unrestricted estimation, for example 3.26E-09, then we would re-estimate the model but in a restricted manner by applying an m x m transition restriction matrix. For all near-zero entries in the transition probability matrix, we placed a zero in the corresponding entry in the transition restriction matrix. The remainder of the entries in the transition restriction matrix were designated NA, indicating that the corresponding transitions probabilities were free to be determined via the restricted estimation. The only caveat of the transition restriction matrix is that each row contains a full set of conditional probabilities, as such each row must sum to one (1) in order for the transition restriction matrix to be correctly specified. Therefore, a row of near-zero values and one non-near-zero value would need to be specified using a one (1) rather than an NA for the non-near-zero entry, since the single non-near-zero entry is no longer free to be estimated, but instead must take the value one (1).

It is very important when using transition restriction matrices, that the coefficients from the original unrestricted estimation are retained as starting values for the following restricted estimation. Otherwise when the restricted estimation is run, the allocation of each volatility regime to a state will be done so in a pseudo-random manner. As such, the restricted transitions identified in the transition restriction matrix will possibly no longer correspond to the correct regime transitions in the subsequent restricted estimation.

3.1 Information criterion

In order to judge the goodness-of-fit of each single-regime model, three information criteria were used, namely: the Bayesian Information Criterion (BIC), the Akaike Information Criterion (AIC) and the Hannan-Quinn Information Criterion (HQC).

$$\begin{aligned}
\text{Bayesian } BIC &= kln\big(n\big) - 2ln\big(\mathcal{L}(\hat{\Theta})\big) \\
\text{Akaike } AIC &= 2k - 2ln\big(\mathcal{L}(\hat{\Theta})\big) \\
\text{Hannan-Quinn } HQC &= -2ln\big(\mathcal{L}(\hat{\Theta})\big) + 2kln\big(ln(n)\big)
\end{aligned}$$

Where Θ is a vector of unknown parameters, $\hat{\Theta}$ is the maximum likelihood estimates of the vector of unknown parameters and k denotes the number of unknown parameters. For reporting the goodness-of-fit in this paper, we negated the information criteria test statistics. As a result, the highest score intuitively corresponds to the optimal model.

The following section presents the estimation results.

4 Results

This section presents the results for the six m-state MRS estimations, with $m \in \{2, ..., 7\}$. For each estimation, we state some or all of the following elements:

- the standard deviations for each state (volatility regime);
- an unrestricted transition matrix;
- a transition restriction matrix;
- a graphical representation of the transition probabilities;
- the probability transition graphs for each state; and finally
- the estimation's goodness-of-fit scores (negated information criteria test statistics).

4.1 2-state MRS estimation results

The standard deviations of the High and Low volatility regimes for the 2-state estimation are listed in Table 4, whilst Table 5 presents the transition probability matrix for the model. The latter table indicates **the presence of volatility clustering in the price data of Bitcoin for the 2-state model**, as demonstrated by the high likelihood for each state to remain in the same state in the following interval. That is, a High (Low) volatility observation is typically followed by a subsequent High (Low) volatility observation. Trivially, there was no requirement for the 2-state model to be re-estimated with a transition restriction matrix applied since none of the transition probabilities had near-zero values.

Table 4. Standard deviations for the volatility regimes (**2-state MRS**)

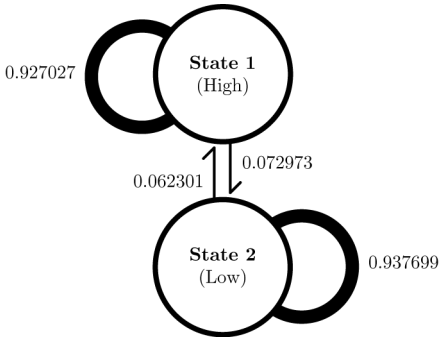
1 (High)	2 (Low)
0.054406	0.016122

Table 5. Unrestricted transition probability matrix (**2-state MRS**)

	1 (High)	2 (Low)
1	0.927027	0.072973
2	0.062301	0.937699

A graphical representation of the transition probabilities presented in Table 5 can be found overleaf in Figure 4. Recall that the thickness of the line for each transition is reflective of the degree of probability for that transition to occur for the following observation.

Figure 4. Unrestricted transition probability diagram (**2-state MRS**)



Figures 5 and 6 illustrate the probability transitions for the 2-state MRS estimation, note the high probability of being in State 1 (High volatility regime) from September 2017 to March 2018, a period which covered the bubble and subsequent crash in the price of Bitcoin.

Figure 5. High state estimation probability transition graph (**2-state MRS**)

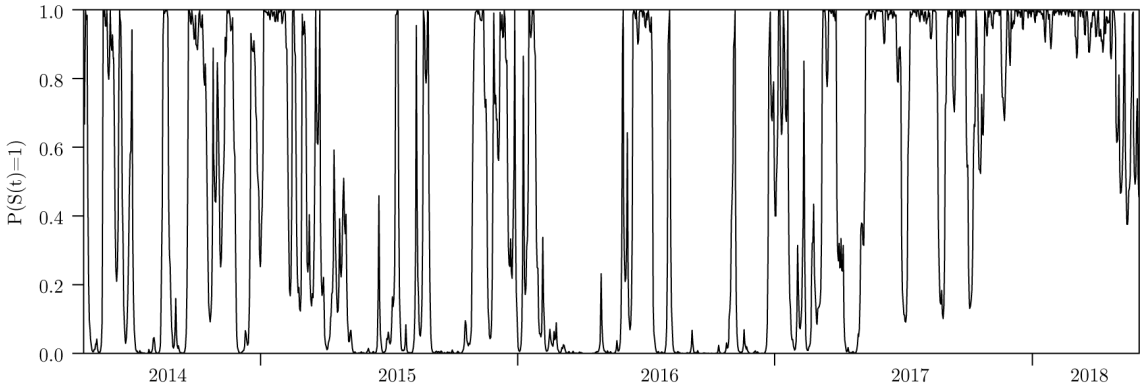


Figure 6. Low state estimation probability transition graph (**2-state MRS**)

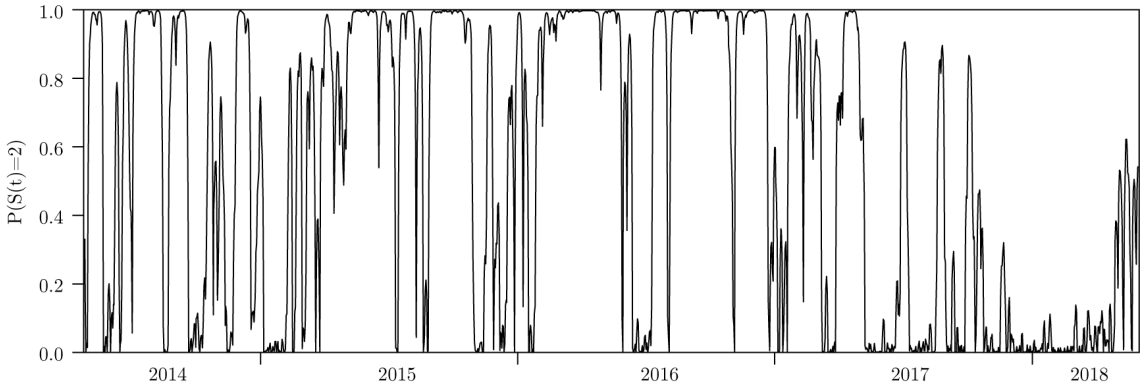


Table 6. Goodness-of-fit scores (**2-state MRS**)

	-AIC	-HQC	-BIC
2-state MRS	4.067042	4.061763	4.052873

The optimal estimation will be selected based upon a goodness-of-fit comparison using three information criteria, namely: Bayesian (BIC); Hannan-Quinn (HQ) and Akaike (AIC). Table 6 presents the goodness-of-fit scores for the 2-state estimation. We have negated the information criteria test statistics; as such the optimal model will be selected based upon the highest goodness-of-fit scores.

4.2 3-state MRS estimation results

The standard deviations of the High, Medium and Low volatility regimes for the 3-state estimation are presented below in Table 7, whilst Table 8 lists the transition probability matrix for the model. Again, **there is evidence of volatility clustering in the price data of Bitcoin**. In addition, the near-zero transition probability from the High state to the Low state in Table 8, indicated that the model would need to be re-estimated using a restrictive transition matrix. As such, a 3 x 3 matrix $\mathbb{A}$ was constructed with the $\mathbb{A}_{13}$ entry set to zero (0). All other transitions would be free to be re-estimated based on the constraint that jumping from the High state to Low state was not permitted.

Table 7. Standard deviations for the volatility regimes (**3-state MRS**)

1 (High)	2 (Medium)	3 (Low)
0.061261	0.026304	0.010667

Table 8. Unrestricted transition probability matrix (**3-state MRS**)

	1 (High)	2 (Medium)	3 (Low)
1	0.923720	0.076280	6.04E-14
2	0.047742	0.906351	0.045907
3	0.015154	0.057529	0.927317

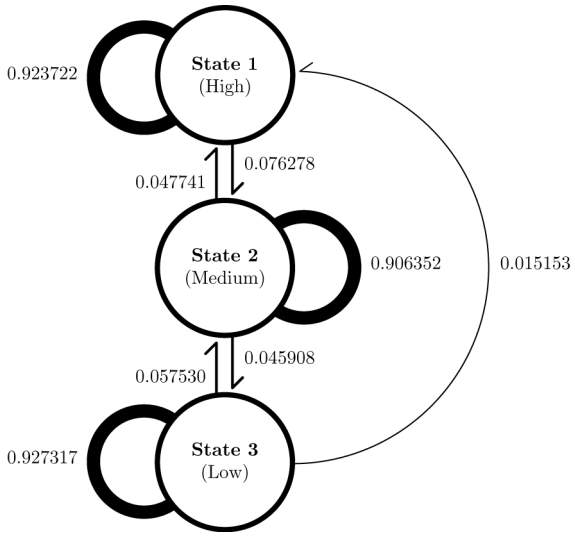
Table 9 (overleaf) states the revised transition probabilities for the restricted 3-state MRS model. As such, setting the probability of transitioning from the High to Low state to zero, resulted in only a marginal change in the remaining unrestricted transition probabilities.

Table 9. Restricted transition probability matrix (**3-state MRS**)

	1 (High)	2 (Medium)	3 (Low)
1	0.923722	0.076278	0
2	0.047741	0.906352	0.045908
3	0.015153	0.057530	0.927317

Figure 7 is a graphical representation of the transition probabilities presented in Table 9. The absence of a transition from the High state to the Low state as a result of applying the transition restriction matrix, is clear to be seen. The diagram indicates that the volatility of Bitcoin can 'jump' from the Low state to the High state, with a probability of 1.5153%. However, for the volatility of Bitcoin to transition from the High regime to the Low regime, it must first pass through the Medium regime in a two-step process, with an overall likelihood of only 0.3502%. Thus, **the sample exhibited an asymmetric tendency in the transition of volatility between different regimes when fitted with the 3-state MRS model**.

Figure 7. Restricted transition probability diagram (**3-state MRS**)



Figures 8–10 illustrate the probability transitions for the 3-state MRS estimation. The most notable feature is the very low probability of transitioning to the Low volatility regime in the most recent 12 months pertaining to the bubble and subsequent crash (Figure 10).

Figure 8. High state estimation probability transition graph (**3-state MRS**)

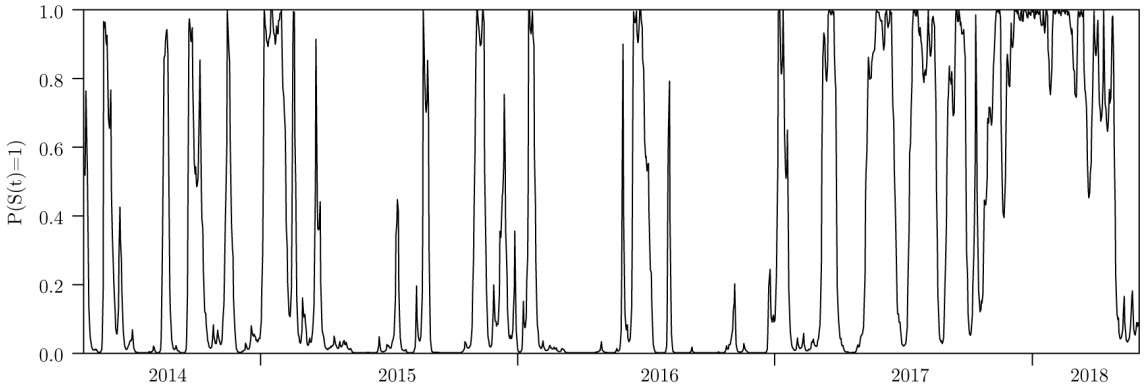


Figure 9. Medium state estimation probability transition graph (**3-state MRS**)

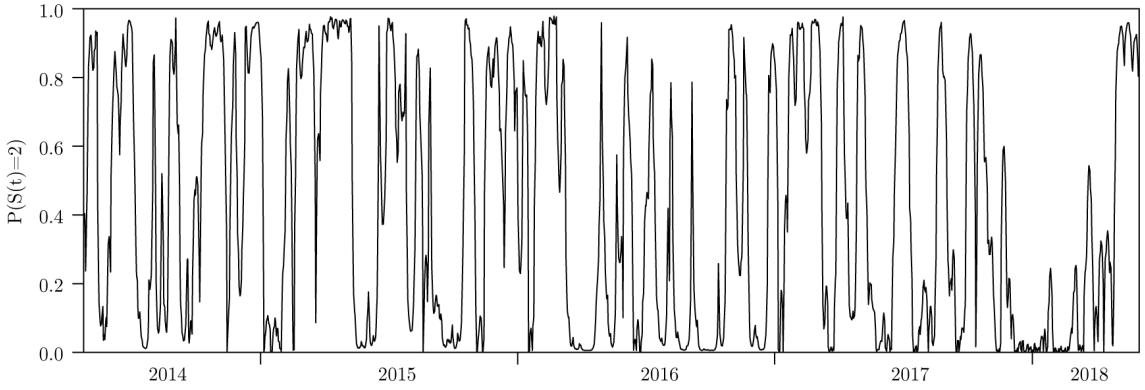


Figure 10. Low state estimation probability transition graph (**3-state MRS**)

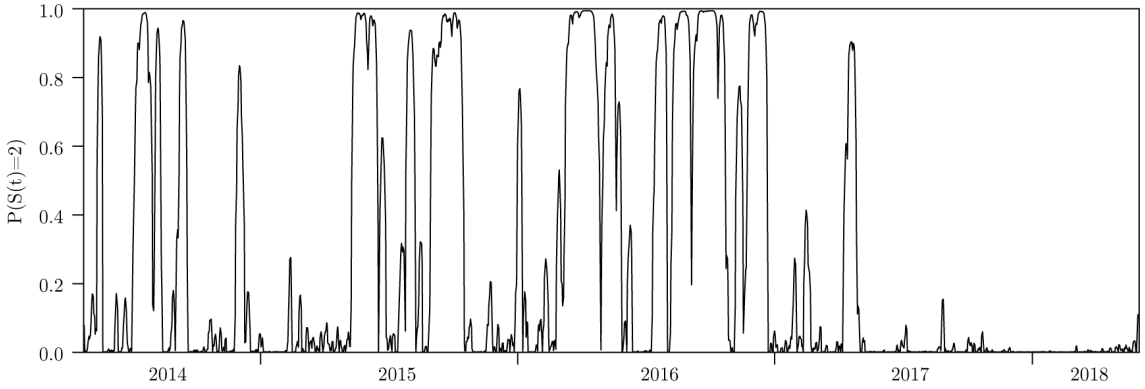


Table 10. Goodness-of-fit scores (**3-state MRS**)

	-AIC	-HQC	-BIC
(unrestricted) 3-state MRS	4.128235	4.116358	4.096355
(restricted) 3-state MRS	**4.129568**	**4.119011**	**4.101231**

Table 10 confirms that the application of a transition restriction matrix to the 3-state MRS estimation was the optimal decision. The resultant goodness-of-fit scores for the restricted estimation were higher than the scores for the unrestricted estimation, for all three negated information criteria.

4.3 4-state MRS estimation results

Tables 11 and 12 respectively present the restricted standard deviations for the volatility regimes and the initial unrestricted transition probabilities estimated with the 4-state MRS model. Due to a number of near-zero values in the transition probability matrix, we constructed a transition restriction matrix $\mathbb{B}$ and re-estimated the 4-state MRS model. See Table 13 (overleaf) for the revised restricted transition probabilities.

Table 11. Restricted standard deviations for the volatility regimes (**4-state MRS**)

1 (High)	2 (Medium)	3 (Low^{+})	4 (Low^{-})
0.063167	0.027787	0.017941	0.006103

Table 12. Unrestricted transition probability matrix (**4-state MRS**)

	1 (High)	2 (Medium)	3 (Low^{+})	4 (Low^{-})
1	0.907850	0.092150	8.1E-129	6.0E-124
2	0.055149	0.914608	0.030243	3.8E-129
3	2.36E-78	4.90E-07	0.499295	0.500705
4	0.026147	0.049639	0.392787	0.531426

$$\mathbb{B} = \begin{bmatrix} NA & NA & 0 & 0 \\ NA & NA & NA & 0 \\ 0 & 0 & NA & NA \\ NA & NA & NA & NA \end{bmatrix}$$

Table 13. Restricted transition probability matrix (**4-state MRS**)

	1 (High)	2 (Medium)	3 (Low$^+$)	4 (Low$^-$)
1	0.907860	0.092140	0	0
2	0.055150	0.914587	0.032631	0
3	0	0	0.499562	0.500438
4	0.026127	0.049749	0.392775	0.531349

As with the 2- and 3-state models, **the 4-state MRS estimation also exhibited volatility clustering**, although to a lesser degree for the lower volatility regimes. In addition, we identified two volatility jumps in the 4-state model pertaining to an upward shock in volatility; both from State 4 to States 1 and 2. In contrast, all of the volatility transition that related to a decrease in volatility were identified solely as transitions between adjacent states. As such, **the 4-state model also demonstrated an asymmetric tendency in the volatility transitions between regimes**, in-line with the findings for the 3-state model. Given the level of persistence in the two higher volatility regimes, any volatility shocks from State 4 could have persisted for a number of subsequent intervals.

Figure 11. Restricted transition probability diagram (**4-state MRS**)

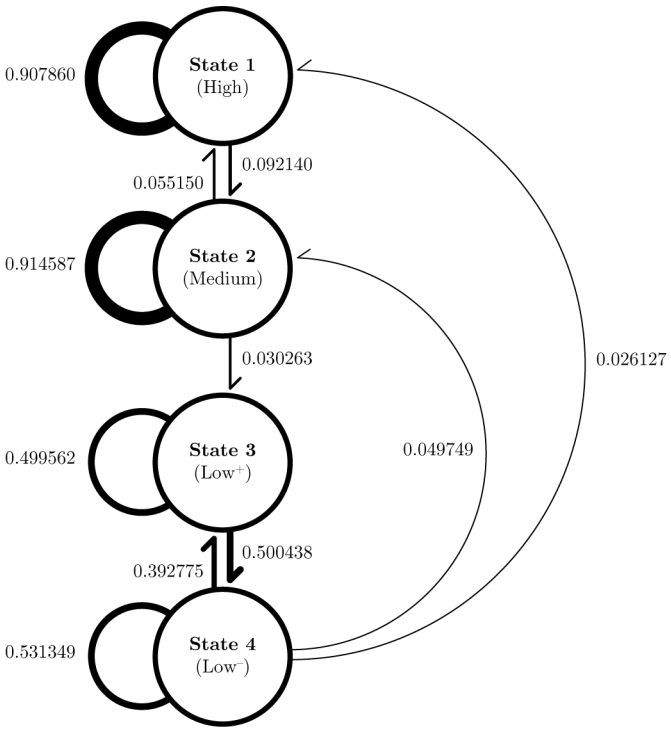


Figure 12. High state estimation probability transition graph (**4-state MRS**)

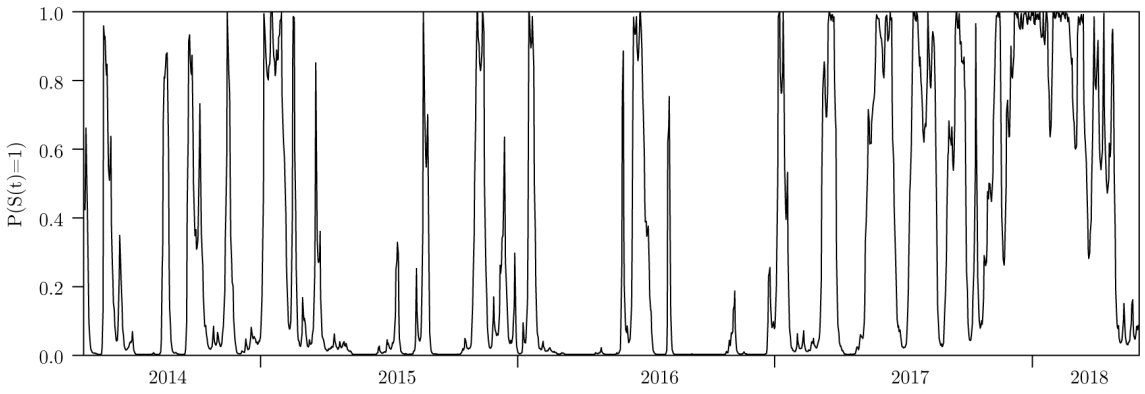


Figure 13. Medium state estimation probability transition graph (**4-state MRS**)

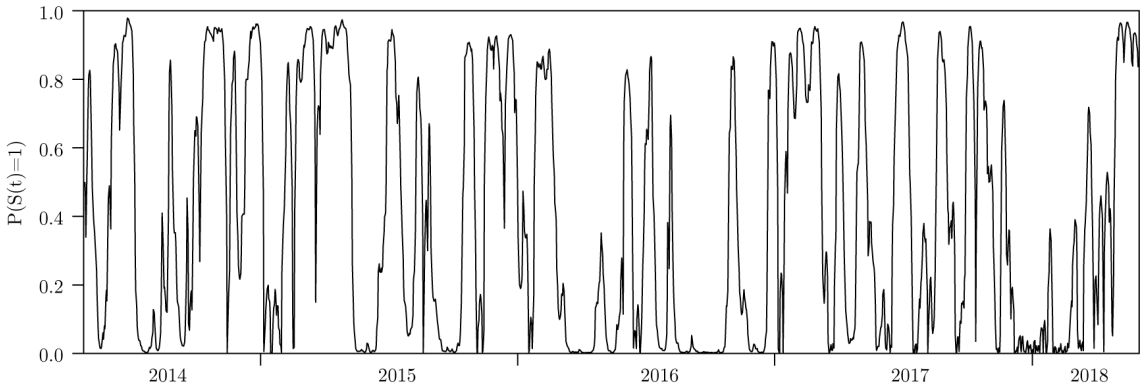


Figure 14. Low^{+} state estimation probability transition graph (**4-state MRS**)

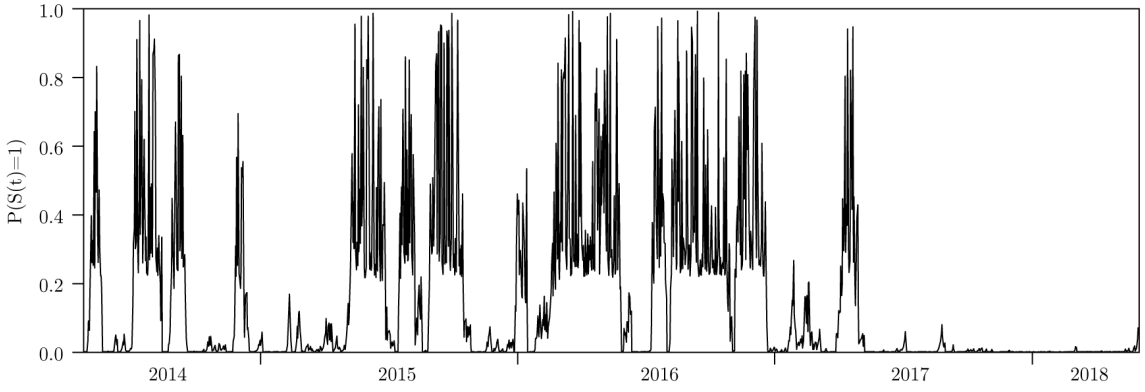
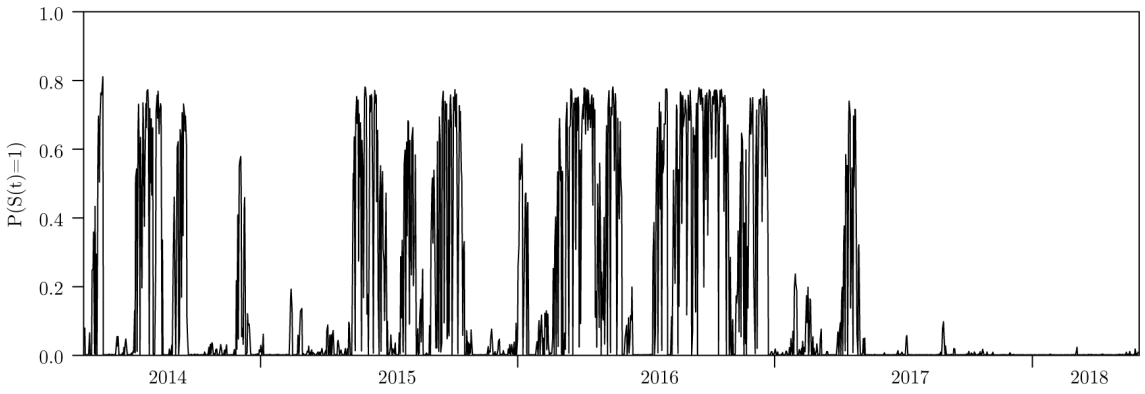


Figure 15. Low^{-} state estimation probability transition graph (**4-state MRS**)



Lastly for the 4-state MRS model, Table 14 presents the goodness-of-fit scores for both runs of the model, namely: the unrestricted and restricted estimations. As can clearly be seen, the unrestricted 4-state estimation generated inferior goodness-of-fit scores as compared to the restricted estimation. As such, the application of a restrictive matrix to the 4-state model was judged to be justified.

Table 14. Goodness-of-fit scores (**4-state MRS**)

	-AIC	-HQC	-BIC
(unrestricted) 4-state MRS	4.144630	4.123516	4.087955
(restricted) 4-state MRS	**4.151296**	**4.136781**	**4.112338**

4.4 5-state MRS estimation results

Table 15 presents the restricted standard deviations for the volatility regimes of the 5-state MRS estimation, whilst Table 16 presents the transition probabilities from the unrestricted estimation of the model.

Table 15. Restricted standard deviations for the volatility regimes (**5-state MRS**)

1 (High)	2 (Medium^{+})	3 (Medium^{-})	4 (Low^{+})	5 (Low^{-})
0.061895	0.041963	0.017825	0.016564	0.005446

Table 16. Unrestricted transition probability matrix (**5-state MRS**)

	1 (High)	2 (Medium^{+})	3 (Medium^{-})	4 (Low^{+})	5 (Low^{-})
1	0.951512	0.048468	1.22E-07	1.94E-05	2.98E-09
2	3.73E-11	0.275730	0.724270	7.09E-09	1.08E-21
3	0.042213	0.370726	0.547947	0.039114	6.23E-26
4	0.010137	0.069137	6.97E-19	0.500808	0.419919
5	3.36E-07	6.89E-06	2.50E-09	0.469091	0.530902

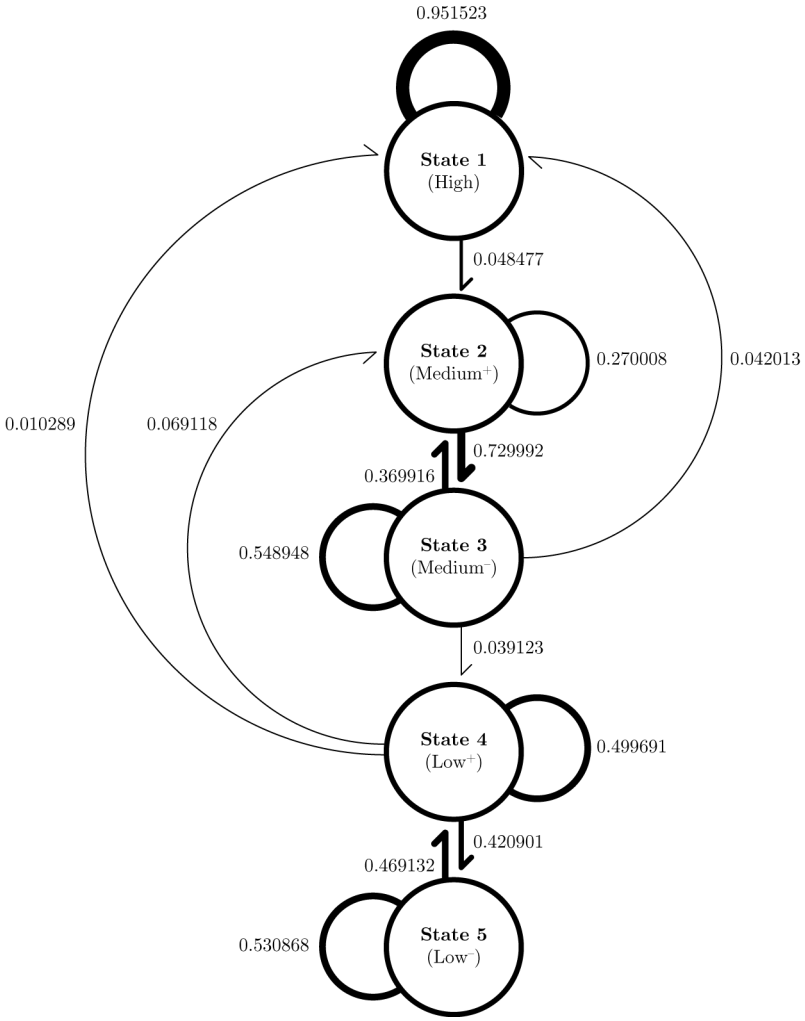
Due to a number of near-zero values in the transition probability matrix presented in Table 16, we constructed the transition restriction matrix $\mathbb{C}$ and re-estimated the 5-state MRS model in a restricted manner. Table 17 lists the resultant restricted transition probabilities, whilst Figure 16 (overleaf) illustrates these restricted transition probabilities.

$$\mathbb{C} = \begin{bmatrix} NA & NA & 0 & 0 & 0 \\ 0 & NA & NA & 0 & 0 \\ NA & NA & NA & NA & 0 \\ NA & NA & 0 & NA & NA \\ 0 & 0 & 0 & NA & NA \end{bmatrix}$$

Table 17. Restricted transition probability matrix (**5-state MRS**)

	1 (High)	2 (Medium^{+})	3 (Medium^{-})	4 (Low^{+})	5 (Low^{-})
1	0.951523	0.048477	0	0	0
2	0	0.270008	0.729992	0	0
3	0.042013	0.369916	0.548948	0.039123	0
4	0.010289	0.069118	0	0.499691	0.420901
5	0	0	0	0.469132	0.530868

Figure 16. Restricted transition probability diagram (**5-state MRS**)



As with the previous estimations of the 2-, 3- and 4-state MRS models, the results of 5-state MRS estimation indicated a degree of persistence in some of the states. Specifically, States 1, 3 and 5 all indicated to varying degrees that an observation in either of these three states would typically remain in that same state in the following interval. The estimation indicated that the High volatility regime had a likelihood of 95.1523% to remain in the High volatility regime in the proceeding interval. As such, **the 5-state MRS model also provided evidence to the presence of volatility clustering in the price data of Bitcoin.**

As with the 2-, 3- and 4-state models, **the 5-state MRS estimation also exhibited volatility clustering**, although to a lesser degree for the lower volatility regimes. In addition, we identified two volatility jumps in the 4-state model pertaining to an upward shock in volatility; both from State 4 to States 1 and 2. In contrast, all of the volatility transition that related to a decrease in volatility were solely identified as transitions between adjacent states. As such, **the 5-state model also demonstrated an asymmetric tendency in the volatility transitions between regimes**, in-line with the findings for the 3-state model. Given the level of persistence in the two higher volatility regimes, any volatility shocks from State 4 could have persisted for a number of subsequent intervals. Figures 17–22 follow on and illustrate the probability transitions for the 5-state MRS estimation.

Figure 17. High state estimation probability transition graph (**5-state MRS**)

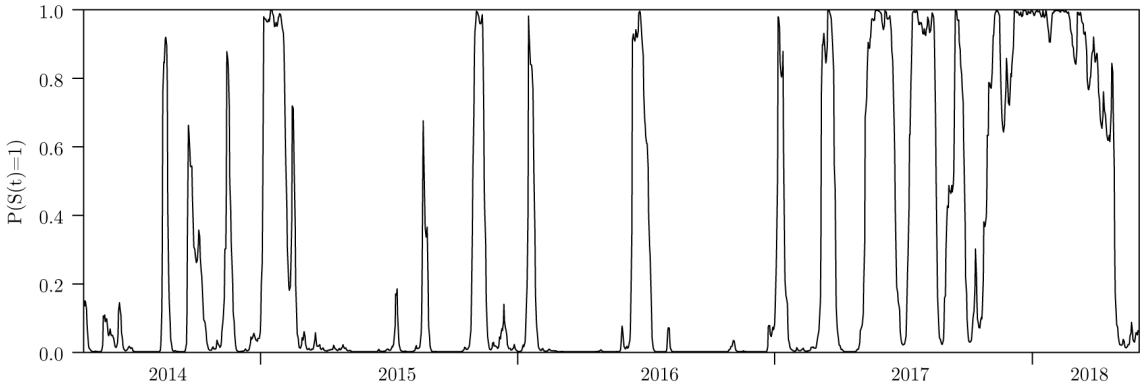


Figure 18. Medium^{+} state estimation probability transition graph (**5-state MRS**)

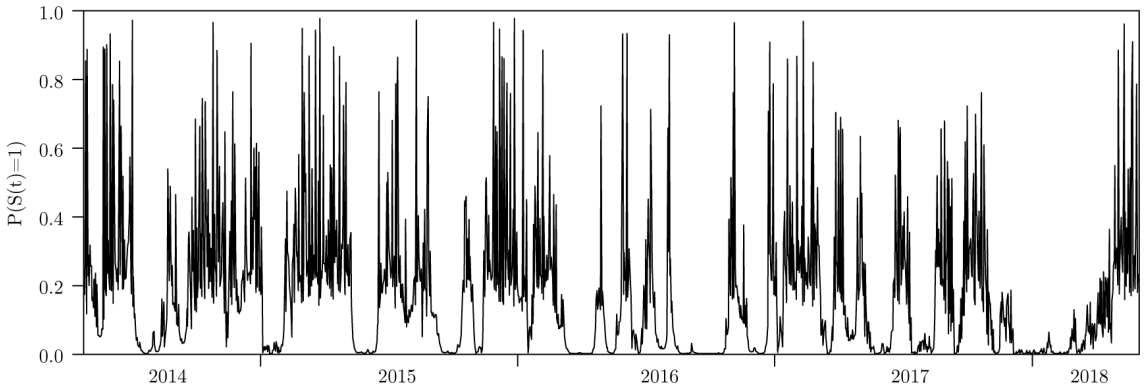


Figure 19. Medium^{-} state estimation probability transition graph (**5-state MRS**)

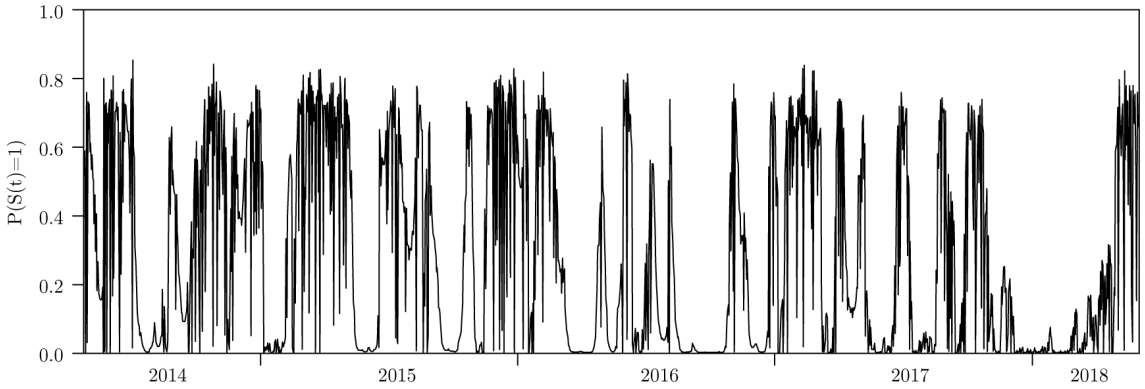


Figure 20. Low^{+} state estimation probability transition graph (**5-state MRS**)

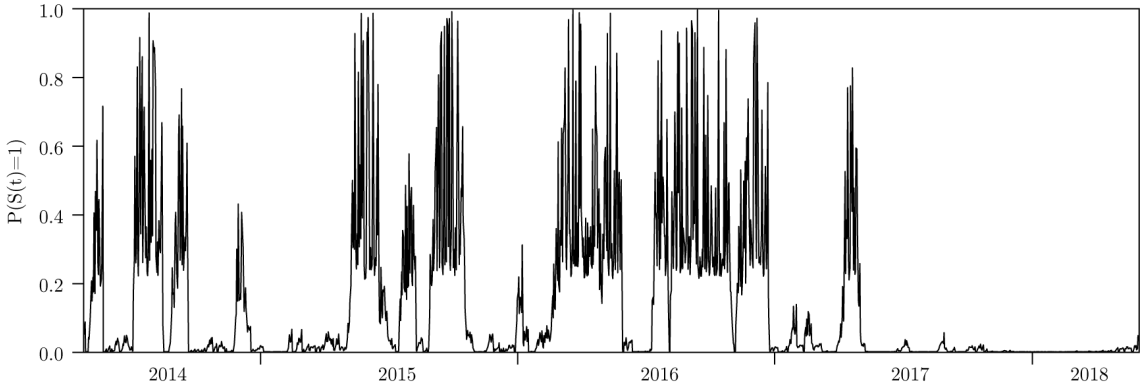
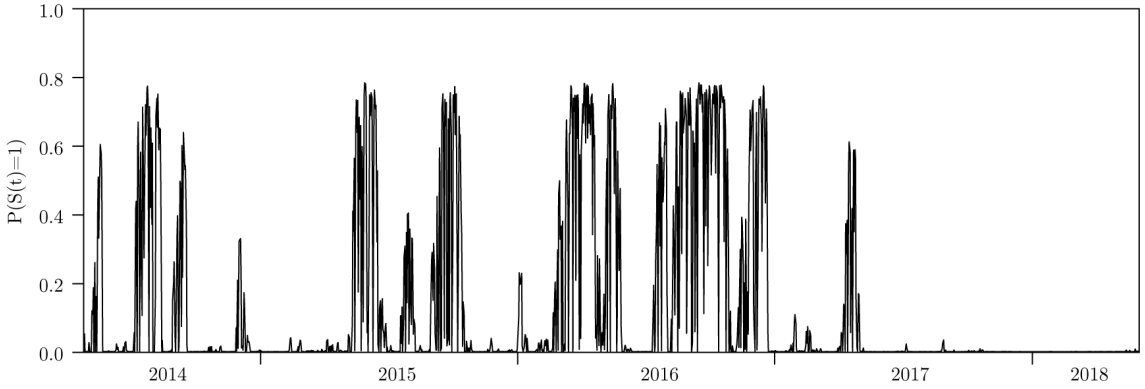


Figure 21. Low^{-} state estimation probability transition graph (**5-state MRS**)



Overall, **the 5-state MRS estimate with a restrictive transition matrix resulted in the highest goodness-of-fit scores for all estimations**. The specific goodness-of-fit scores for the 5-state model are presented in Table 18 (below).

Table 18. Goodness-of-fit scores (**5-state MRS**)

	-AIC	-HQC	-BIC
(unrestricted) 5-state MRS	4.149319	4.116330	4.060765
(restricted) 5-state MRS	**4.163990**	**4.145515**	**4.114400**

4.5 Overfitted MRS estimation results

In light of a degradation in the unrestricted goodness-of-fit scores for the 6- and 7-state models with respect to the 5-state model's unrestricted estimation; coupled with the presence of absorbing states in the unrestricted estimation transition probabilities for the 6- and 7-state models; and the nonsensical probability transitions for the 7-state model, we determined that both the 6- and 7-state unrestricted estimations were overfitting the sample. As such, continuing to then fit a transition restriction matrix and re-estimate the sample would only exacerbate the issues stated. Therefore, we discounted the 6- and 7-state models from any further consideration as optimal MRS models for capturing the regime heteroskedastoicity of Bitcoin. The following two subsections present the results of the unrestricted estimations for the 6- and 7-state models for completeness only.

4.5.1 6-state MRS estimation results

Tables 19 and 20 present the unrestricted standard deviations and transition probabilities for the 6-state model. As can be seen in Table 20 (overleaf), State 5 was an absorbing state for State 3, in that all transitions out of State 3 only transitioned to State 5. In addition, State 4 was an absorbing state for State 1. The presence of absorbing states in the estimation of an MRS model for a sufficiently large, non-linear time series indicated overfitting of the sample. Figures 22 and 23 (overleaf) present the probability transitions for the 6-state model's unrestricted estimation and Figure 24 illustrates the unrestricted transition probabilities for the 6-state model.

Table 19. Unrestricted standard deviations for the volatility regimes (**6-state MRS**)

1 (High^{+})	2 (High^{-})	3 (Med^{+})	4 (Med^{-})	5 (Low^{+})	6 (Low^{-})
0.067075	0.063763	0.028083	0.026809	0.012592	0.005661

Table 20. Unrestricted transition probability matrix (**6-state MRS**)

	1 (High$^+$)	2 (High$^-$)	3 (Med$^+$)	4 (Med$^-$)	5 (Low$^+$)	6 (Low$^-$)
1	≈ 0	≈ 0	≈ 0	1.000000	≈ 0	≈ 0
2	≈ 0	0.940486	0.059335	≈ 0	≈ 0	≈ 0
3	≈ 0	≈ 0	≈ 0	≈ 0	1.000000	≈ 0
4	0.098229	0.036291	0.019440	0.831164	0.010623	0.004253
5	≈ 0	≈ 0	≈ 0	≈ 0	0.706646	0.293335
6	0.010289	≈ 0	0.331971	0.042013	≈ 0	0.620334

Figure 22. Probability transition graphs for States 1, 2 and 3 (**6-state MRS**)

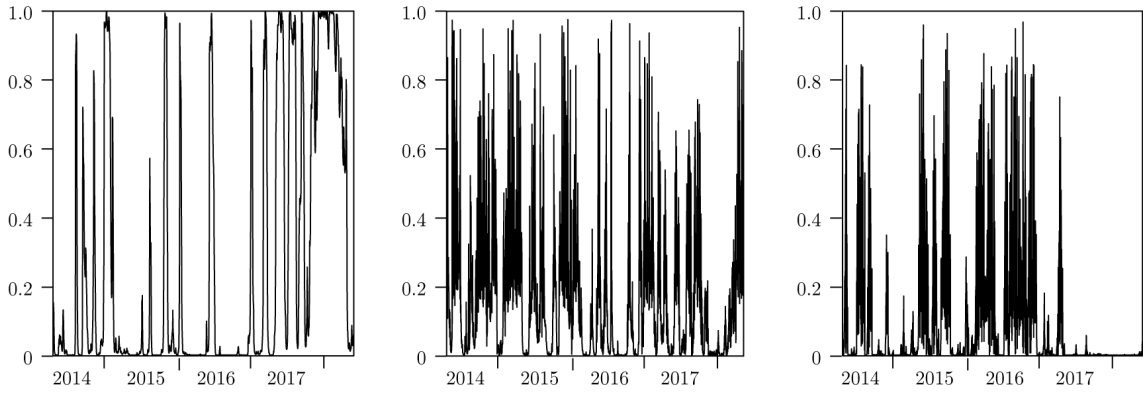


Figure 23. Probability transition graphs for States 4, 5 and 6 (**6-state MRS**)

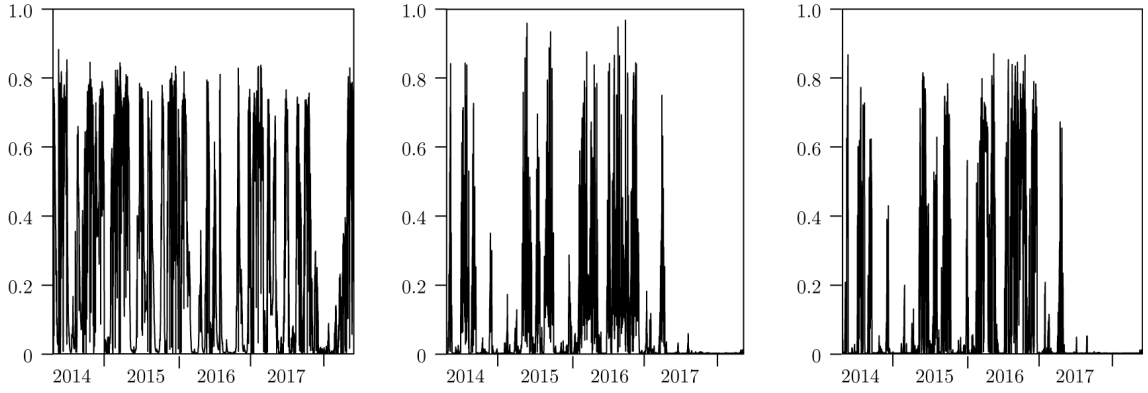


Figure 24. Unrestricted transition probability diagram (**6-state MRS**)

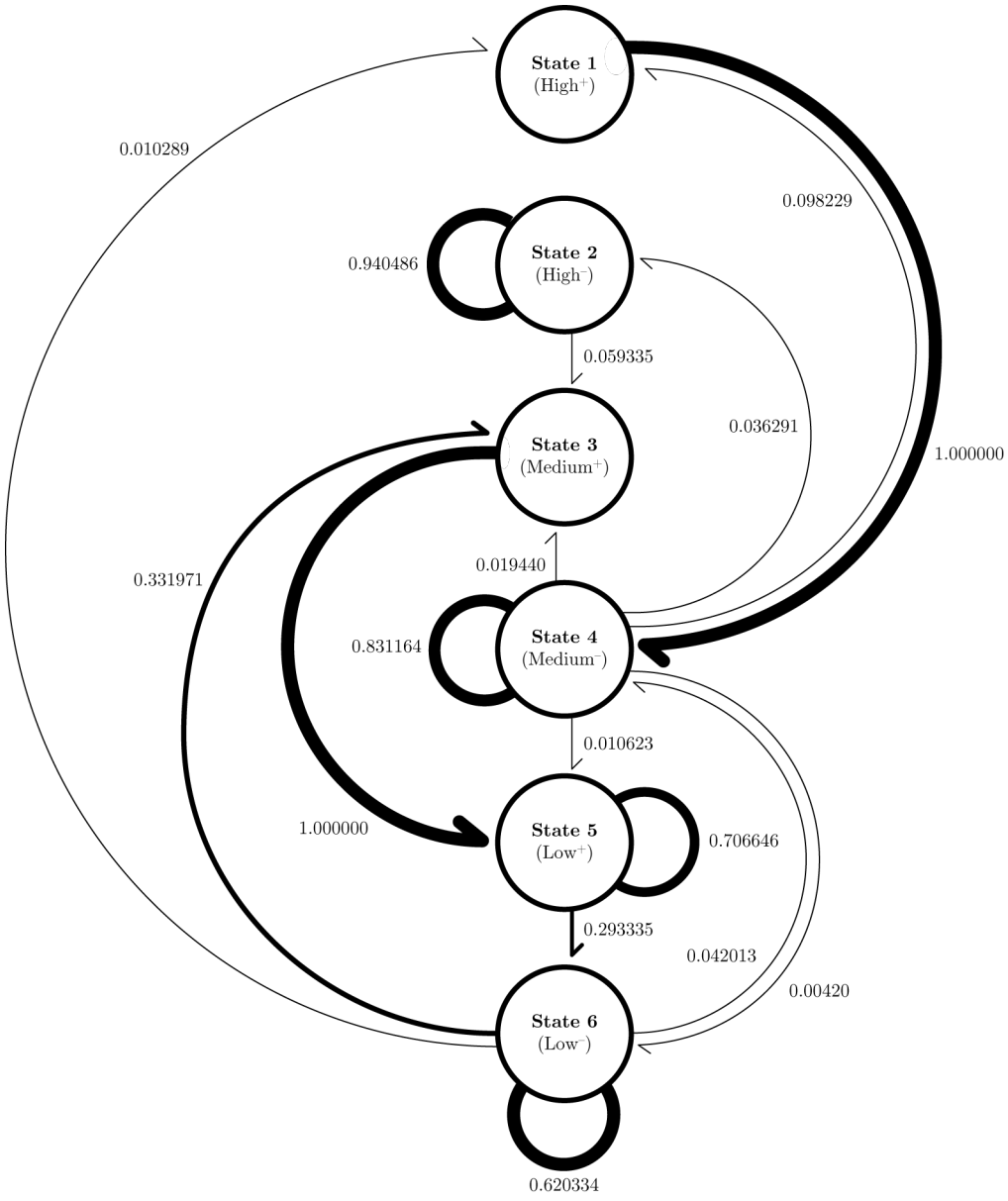


Table 21 presents the goodness-of-fit scores for the unrestricted 6-state MRS estimation.

Table 21. Goodness-of-fit scores (**6-state MRS**)

	-AIC	-HQC	-BIC
(unrestricted) 6-state MRS	4.130433	4.082928	4.002916

4.5.2 7-state MRS estimation results

Tables 22 and 23 present the unrestricted standard deviations and transition probabilities for the 7-state MRS estimation respectively. Evidence of the sample being overfitted was immediately apparent in the transition probabilities. As can clearly be seen in the last column of Table 23, none of the volatility regimes in the 7-state estimation transitioned to State 7, not even State 7 itself. In addition, Figures 25–27 (overleaf) illustrate the probability transitions for the 7-state model. Note the State 7 (Low^{-}) probability transitions are fixed at zero (0) for the sample window, confirmation of the data being overfitted.

Table 22. Unrestricted standard deviations for the volatility regimes (**7-state MRS**)

1 (High^{+})	2 (High^{-})	3 (Med^{+})	4 (Med)	5 (Med^{-})	6 (Low^{+})	7 (Low^{-})
0.061895	0.041963	0.017825	0.016564	0.005446	0.005446	0.005446

Table 23. Unrestricted transition probability matrix (**7-state MRS**)

	1 (High^{+})	2 (High^{-})	3 (Med^{+})	4 (Med)	5 (Med^{-})	6 (Low^{+})	7 (Low^{-})
1	0.950191	0.049305	≈ 0	≈ 0	≈ 0	≈ 0	≈ 0
2	≈ 0	≈ 0	≈ 0	1.000000	≈ 0	≈ 0	≈ 0
3	0.019025	0.290815	0.536908	≈ 0	0.153255	≈ 0	≈ 0
4	≈ 0	≈ 0	≈ 0	0.550525	≈ 0	0.444848	≈ 0
5	0.157049	0.628382	≈ 0	0.214568	≈ 0	≈ 0	≈ 0
6	≈ 0	≈ 0	≈ 0	0.444848	≈ 0	0.550907	≈ 0
7	≈ 0	0.318544	≈ 0	0.344241	0.336985	≈ 0	≈ 0

Figure 28 illustrates the transition probability matrix for the 7-state model. As can be seen in Table 23 and Figure 28, State 4 was an absorbing state for State 2. The goodness-of-fit scores for the model were as follows: -AIC 4.122400; -HQC 4.057740; and -BIC 3.948834.

Figure 25. Probability transition graphs for States 1, 2 and 3 (**7-state MRS**)

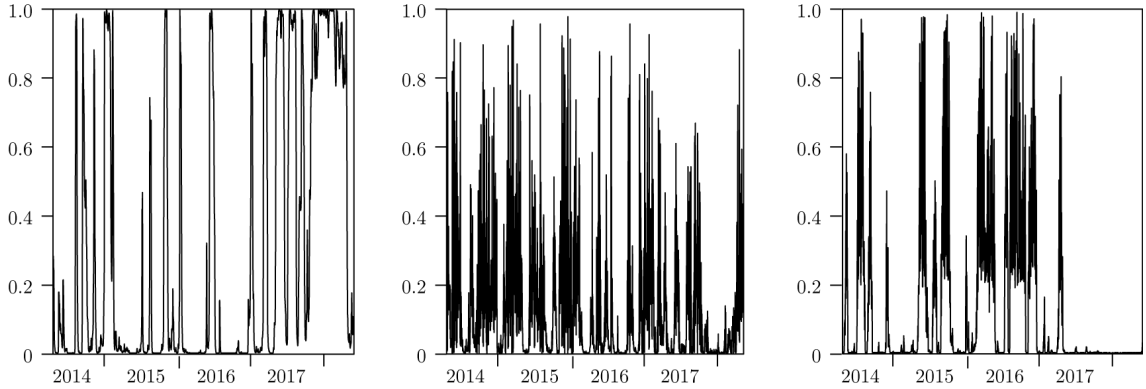


Figure 26. Probability transition graphs for States 4, 5 and 6 (**7-state MRS**)

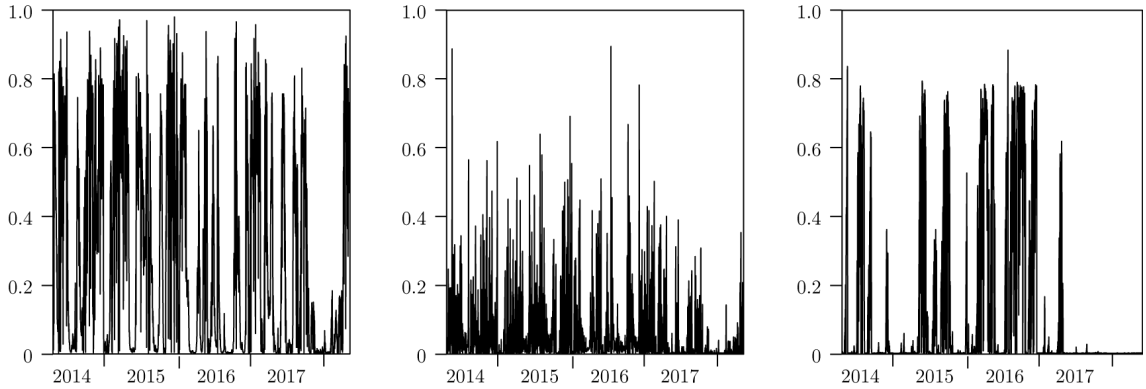


Figure 27. Probability transition graph for State 7 (**7-state MRS**)

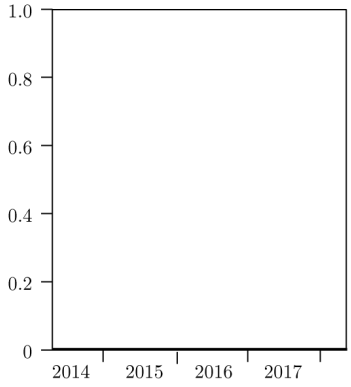
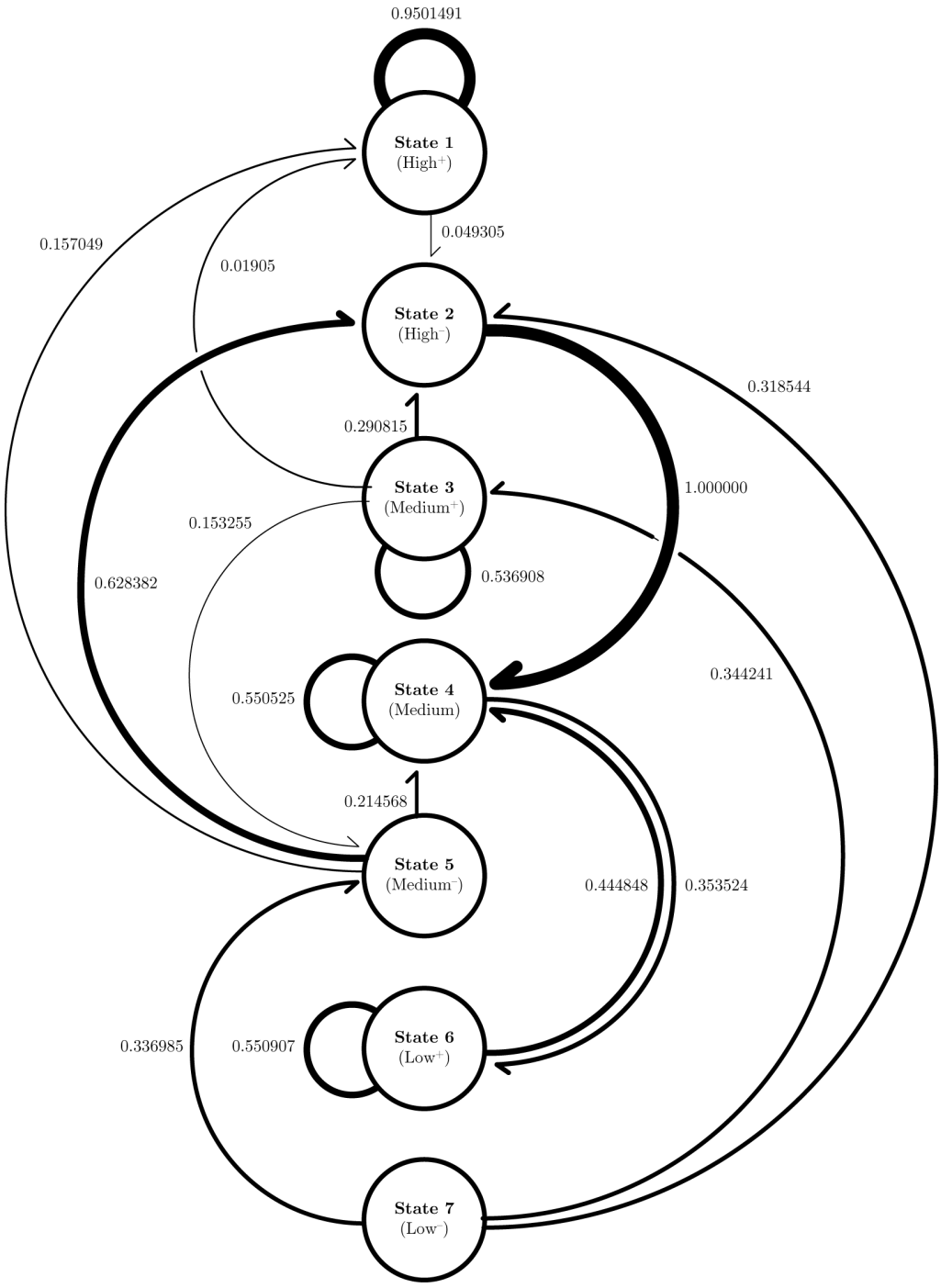


Figure 28. Unrestricted transition probability diagram (**7-state MRS**)



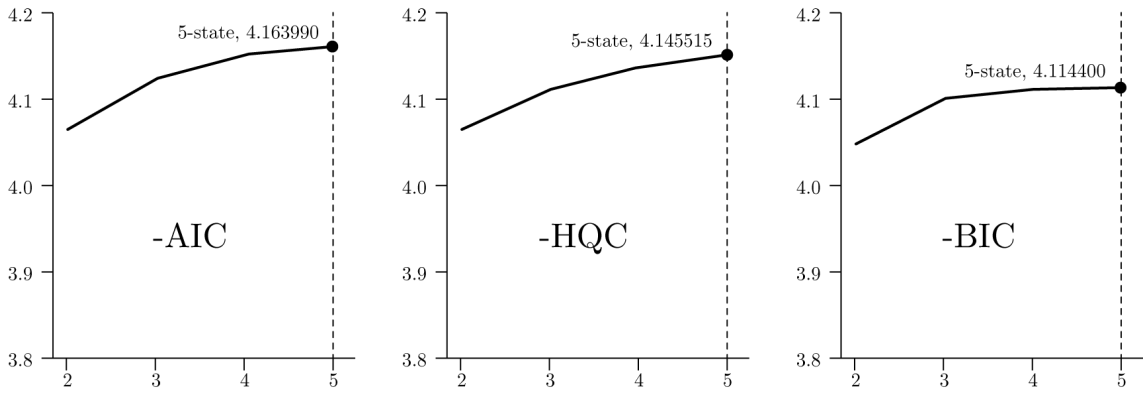
4.6 Goodness-of-fit results

Table 24 presents the amalgamated goodness-of-fit scores for the 2-state model and the three completed restricted estimations, i.e. 3-, 4-, and 5-state models. As such, **the optimal model, as judged by the goodness-of-fit criteria, is unanimously the restricted 5-state model**.

Table 24. Restricted goodness-of-fit scores (***m*-state MRS, with $m \in \{2, ..., 5\}$**)

	-AIC	-HQC	-BIC
2-state MRS	4.067042	4.061763	4.052873
(restricted) 3-state MRS	4.129568	4.119011	4.101231
(restricted) 4-state MRS	4.151296	4.136781	4.112338
(restricted) 5-state MRS	**4.163990**	**4.145515**	**4.114400**

Figure 29. Restricted goodness-of-fit scores versus number of states (**-AIC, -HQC, -BIC**)

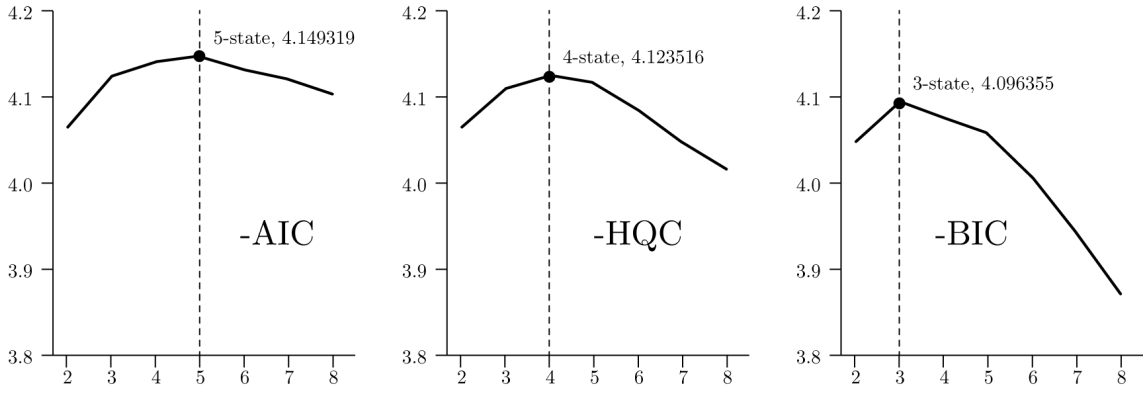


For completeness, Table 25 (overleaf) lists the goodness-of-fit scores for all of the unrestricted estimations along with the 2-state model for comparative purposes. In addition, we included the goodness-of-fit scores for the previously unmentioned 8-state model. This model was included to confirm the downward trend in the unrestricted goodness-of-fit scores as the estimations became overfitted for the higher state models (6- and 7-state versions). For all of the models estimated in Table 25, only the 2-state model's scores should be considered as reflective of the optimal estimation for that model. Since the 3- to 8-state models all required re-estimating using a transition restriction matrix. When comparing like-for-like results between Tables 24 and 25, it is clear to see that the restricted estimations outperformed the unrestricted estimations for all of the comparative models.

Table 25. Unrestricted goodness-of-fit scores (***m*-state MRS, with $m \in \{2, ..., 8\}$**)

	-AIC	-HQC	-BIC
2-state MRS	4.067042	4.061763	4.052873
(unrestricted) 3-state MRS	4.128235	4.116358	**4.096355**
(unrestricted) 4-state MRS	4.144630	**4.123516**	4.087955
(unrestricted) 5-state MRS	**4.149319**	4.116330	4.060765
(unrestricted) 6-state MRS	4.130433	4.082928	4.002916
(unrestricted) 7-state MRS	4.122400	4.057740	3.948834
(unrestricted) 8-state MRS	4.102808	4.018355	3.876111

Figure 30. Unrestricted goodness-of-fit scores versus number of states (**-AIC, -HQC, -BIC**)



4.7 Estimation runtimes

We found that as the complexity of the model increased, the runtime required to complete each estimation significantly increased in duration. This was due to the exponential increase in the number of operations that needed to be completed in order to account for the increased dimensionality due to each additional state. Table 26 (overleaf) presents the mean score of five unrestricted estimations per m-state model, with $m \in \{2, ..., 8\}$. Whilst the 2-state MRS estimation took only 90 seconds to complete, the 7- and 8-state MRS estimations took well over an hour on average and almost two hours for the latter estimation. In addition, it should be noted that the runtimes listed in Table 26 are only for the initial unrestricted estimation. As such, the times presented do not account for the determination and construction of a transition restriction matrix and the re-estimation for

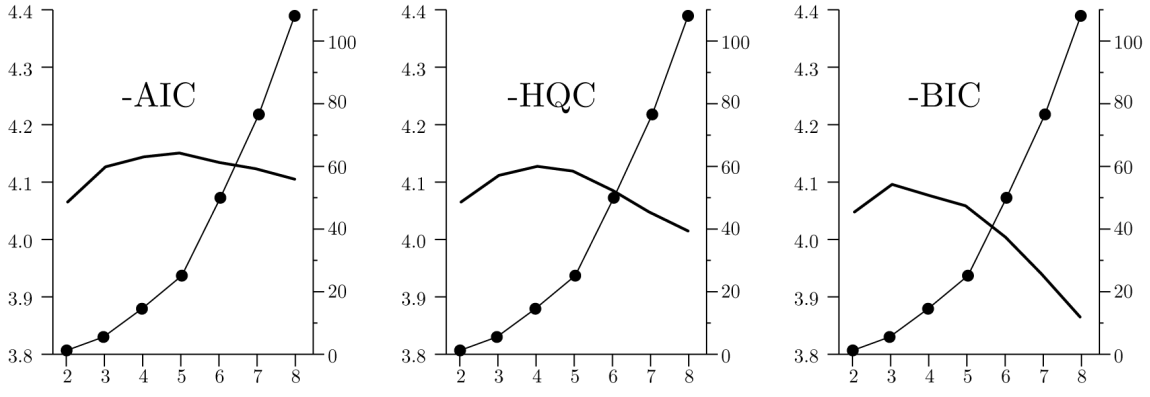
the sample using a restricted model for the 3-state or higher models. Whilst the complexity of the restricted estimation is lower than that of the unrestricted estimation, due to the fewer number of operations that need to be computed, we found the reduction in runtime was only between $10 - 20\%$ that of the original unrestricted estimation.

In making the case for an optimal model, it could be argued that the difference in the goodness-of-fit scores for the restricted 3-, 4- and 5-state estimations with respect to the -BIC is negligible. However, the difference in runtime between the three models is sizable in the context a 500-interval rolling forecast exercise. As such, the determination of an optimal model independent of the runtime required to complete the necessary estimation is myopic thinking. The runtimes presented in this paper, however were achieved using a standard laptop processor working in isolation. A network of more powerful workstations operating in unison would be able to complete said exercise in a much shorter timeframe.

Table 26. Runtimes (***m*-state MRS, with $m \in \{2, ..., 8\}$**)

	Runtime (hr:min:sec)
2-state MRS	00:01:27.69
(unrestricted) 3-state MRS	00:05:12.74
(unrestricted) 4-state MRS	00:13:00.24
(unrestricted) 5-state MRS	00:25:29.11
(unrestricted) 6-state MRS	00:50:25.77
(unrestricted) 7-state MRS	01:27:11.08
(unrestricted) 8-state MRS	01:49:32.38

Figure 31. Unrestricted goodness-of-fit scores versus runtime (minutes) (**-AIC, -HQC, -BIC**)



4.8 Persistence of volatility shocks

Motivated by the volatility jumps illustrated in the transition probability diagrams for the 3-, 4- and 5-state restricted estimations, we reviewed the corresponding transition restriction matrices:

$$\mathbb{A} = \begin{bmatrix} NA & NA & 0 \\ NA & NA & NA \\ NA & NA & NA \end{bmatrix}$$

$$\mathbb{B} = \begin{bmatrix} NA & NA & 0 & 0 \\ NA & NA & NA & 0 \\ 0 & 0 & NA & NA \\ NA & NA & NA & NA \end{bmatrix}$$

$$\mathbb{C} = \begin{bmatrix} NA & NA & 0 & 0 & 0 \\ 0 & NA & NA & 0 & 0 \\ NA & NA & NA & NA & 0 \\ NA & NA & 0 & NA & NA \\ 0 & 0 & 0 & NA & NA \end{bmatrix}$$

Each of the three transition restriction matrices contained triangles of zero entries in their off-diagonal upper corners, although only trivially in the case of the 3 x 3 matrix $\mathbb{A}$. The position of these zeroes indicated that transitions from higher volatility regimes to lower volatility regimes tended to occur in an ordered manner without any jumps. However, the presence of NA entries in the off-diagonal lower corners of these matrices, indicated that volatility could not only increase in a sequentially ordered manner, but that volatility could also jump transition from the lowest regimes to the highest regimes. The following matrix entries correspond to volatility jumps in the above matrices: $\mathbb{A}_{31}$; $\mathbb{B}_{41}$; $\mathbb{B}_{42}$; $\mathbb{C}_{31}$; $\mathbb{C}_{41}$; and $\mathbb{C}_{42}$. As such volatility shocks for Bitcoin, in the form of a jump from a lower regime to a higher regime in a single transition, would require more than a single subsequent interval to correct. In modelling the regime heteroskedasticity of the sample **using applying Markov regime-switching models, we have shown there exists a persistence associated with the volatility shocks in the price data of Bitcoin.**

5 Conclusion

In light of Molnár and Thies (2018) demonstrating that the price data of Bitcoin contained seven distinct volatility regimes, and in response to the recent paper by Ardia et al (2018), who only fitted the 2- and 3-state MRS-GRACH models to the price data of Bitcoin; we fitted a sample of Bitcoin returns with six m-state MRS estimations, with $m \in \{2, ..., 7\}$. Our aim was to identify the optimal number of states for modelling the regime heteroskedasticity in the price data of Bitcoin. In doing so, we found that **the restricted 5-state Markov regime-switching model attained the highest goodness-of-fit scores** in our comparative study. However, for each additional state over the simple 2-state model that was estimated, there was an increased complexity in the form of transition restriction matrices and a disproportionate marginal cost in the form of computational runtime. Whilst we did attempt to fit both the 6- and 7-state models to our sample, in reference to Molnár and Thies' assertion; we found that the estimation results indicated overfitting of the sample, in the form of absorbing states and a redundant regime (State 7).

By applying Markov regime-switching models to the sample, we also found evidence of:

- **volatility clustering**, high degree of persistence in the 2-, 3- and 4-state estimations;
- **volatility jumps**, non-sequential transitions in the 3-, 4- and 5-state estimations;
- **asymmetric volatility transitions**, presence of volatility steps and jumps for increases in volatility, but only volatility steps for decreases in volatility; and
- **shock persistence**, presence of asymmetric volatility transitions, so that upward shocks in the volatility of Bitcoin typically persisted beyond a single interval.

The estimation of conditional heteroskedasticity in the time series of Bitcoin returns without any consideration for the evident regime heteroskedasticity, is certainly a fool's errand. As such, future research should consider extending Ardia et al.'s methodology to include 4- and 5-state MRS-GARCH versions for the modelling of the cryptocurrency's price data. In addition, the use of GARCH variants such as EGARCH and TGARCH should also be considered within the framework of modelling regime heteroskedasticity using the Markov switching methodology, i.e. MRS-EGARCH and MRS-TGARCH.

Cryptocurrencies are being viewed more and more as a store of value and not just by a few feverish converts on a cryptocurrency message board. Increasing interest in cryptocurrencies necessitates a better understanding of the volatility dynamics of these instruments and the pursuit of considered and robust risk management tools. As such, the results contained within this paper will be useful to cryptocurrency stakeholders from an option pricing and risk management perspective.

References

Alizadeh, A., Nomikos, N. and Pouliasis, P. (2008) A Markov regime switching approach for hedging energy commodities. *Journal of Banking and Finance* **32**, 1970–1983.

Ardia, D., Bluteau, K., Boudt, K. and Catania, L. (2017) Forecasting risk with Markov-switching GARCH models: A large-scale performance study. *International Journal of Forecasting*, **34**(40, 733-747.

Ardia, D., Bluteau, K. and Rüede, M. (2018) Regime changes in Bitcoin GARCH volatility dynamics. *Finance Research Letters*. Available online: 10th August 2018.

Ang, A. and Timmermann, A. (2012) Regime Changes and Financial Markets. *Annual Review of Financial Economics*, **4**, 313–337.

Baek, C. and Elbeck, M. (2015) Bitcoins as an investment or speculative vehicle? A first look. *Applied Economic Letters*, **22**(1), 30–34.

Bariviera, A. (2017) The inefficiency of Bitcoin revisited: A dynamic approach. *Economic Letters*, **161** 1–4.

Bariviera, A., Basgall, M., Hasperue, W. and Naiouf, M. (2017) Some stylized facts of the Bitcoin market. *Physica A: Statistical Mechanics and its Applications*, **484**, 82–90.

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

Baur, D., Dimpfl, T. and Kuck, K. (2017) Bitcoin, gold and the US dollar A replication and extension. *Finance Research Letters*.

Baur, D., Hong, K. and Lee, A. (2017) Bitcoin: Medium of exchange or speculative assets? *Journal of International Financial Markets, Institutions and Money*.

Bauwens, L., Backer, B. and Dufays, A. (2014) A Bayesian method of change-point estimation with recurrent regimes: Application to GARCH models. *Journal of Empirical Finance*, **29**, 207–229.

Bauwens, L., Preminger, A. and Rombouts, J. (2010) Theory and inference for a Markov switch GARCH model. *Econometrics Journal*, **13**, 218–244.

Berndt, E. (1991) *The Practice of Econometrics*, Reading, MA, Addison-Wesley.

Black, F. (1976) Studies of stock price volatility changes. In: Proceedings of the 1976 Meetings of the American Statistical Association, 171–181.

Blau, B. (2017) Price dynamics and speculative trading in Bitcoin. *Research in International Business and Finance*, **41**, 493–499.

Bollerslev, T. (1986) Generalised autoregressive conditional heteroskedasticity. *Journal of Econometrics*, **31**, 307–327.

Bouoiyour, J. and Selmi, R. (2015) Bitcoin price: Is it really that new round of volatility can be on way? *Munich Personal RePEc Archive*, **6558**.

Cai, J. (1994) A Markov model of unconditional variance in ARCH. *Journal of Business and Economic Statistics*, **12**, 309–316.

Campbell, J., Lo, A. and Mackinlay, A. (1997) *The Econometrics of Financial Markets*, Princeton, NJ, Princeton University Press.

Chang,Y., Choi, Y and Park, J. (2017) A new approach to model regime switching. *Journal of Econometrics*, **196**, 127–143.

Cheah, E. and Fry, J. (2015) Speculative bubbles in Bitcoin markets? An empirical investigation into the fundamental value of Bitcoin. *Economic Letters*, **130**, 32–36.

Chu, J., Chan, S., Nadarajah, S. and Osterrieder, J. (2017) GARCH Modelling of Cryptocurrencies. *Journal of Risk and Financial Management*, **10**, 17.

Ding, Z., Granger, C. and Engle, R. (1993) A long memory property of stock market returns and a new model. *Journal of Empirical Finance*, **1**, 83–106.

Dyhrberg, A. (2016a) Bitcoin, gold and the dollar A GARCH volatility analysis. *Finance Research Letters*, **16**, 85–92.

Dyhrberg, A. (2016b) Hedging capabilities of Bitcoin. Is it the virtual gold? *Finance Research Letters*, **16**, 139–144.

Elendner, H., Trimborn, S., Ong, B. and Lee, T. (2018) Investing in Crypto-Currencies Beyond Bitcoin. *Handbook of Blockchain, Digital Finance, and Inclusion*, **1**, 145–173.

Enders, W. (2001) *Applied Econometric Time Series*, 2nd ed., Hoboken, NJ, Wiley.

Engle, R. (1982) Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, **50**, 987–1007.

Engle, R. and Bollerslev, T. (1986) Modelling the persistence of conditional variances. *Econometric Reviews*, **5**, 1–50.

Engle, R., Lilien, D. and Robins, R. (1987) Estimating time varying risk premia in the term structure: the ARCH-M model. *Econometrica*, **55**, 391–407.

Ghahramani, Z. (2001) An Introduction to Hidden Markov Models and Bayesian Networks. *International Journal of Pattern Recognition and Artifcial Intelligence*, **15**(1), 9–42.

Glosten, L., Jagannathan, R. and Runkle, D. (1993) On the relation between the expected value and the volatility of the nominal excess return on stocks. *Journal of Finance*, **48**, 1779–1801.

Hamilton, J. (1989) A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle. *Econometrica*, **57**, 357–384.

Hamilton, J. (1994) *Time Series Analysis*, Princeton, NJ, Princeton University Press.

Hamilton, J. and Susmel, R. (1994) Autoregressive conditional heteroskedasticity and changes in regime. *Journal of Econometrics*, **64**, 307–333.

Harvey, A. (1993) *Time Series Models*, 2nd ed., Harlow, FT/Prentice Hall.

Haas, M., Mittnik, S. and Paolella, M. (2004) A new approach to Markov-switching GARCH models. *Journal of Financial Econometrics*, **2**, 493–530.

Henry, O. (2009) Regime switching in the relationship between equity returns and short-term interest rates in the UK. *Journal of Banking and Finance* **33**, 405–414.

Hentschel, L. (1995) All in the family: nesting symmetric and asymmetric GARCH models. *Journal of Financial Economics*, **39**, 71–104.

Hubbert, S. (2012) *Essential mathematics for market risk management*. 2nd ed., Chichester, John Wiley & Sons Ltd.

Iqbal, F. (2016) Forecasting Volatility and Value-at-Risk of Pakistan Stock Market with Markov Regime-Switching GARCH Models. *European Online Journal of Natural and Social Sciences*, **5**(1), 172–189.

Jiang, Y., Nie, H. and Ruan, W. (2017 Time-varying long-term memory in Bitcoin market. *Finance Research Letters*.

Johnston, J. and DiNardo, J. (1997) *Econometric Methods*, 4th ed., London, McGraw Hill.

Katsiampa, P. (2017) Volatility estimation for Bitcoin: A comparison of GARCH models. *Economics Letters*, **158**, 3–6.

Kim, C. (1994) Dynamic Linear Models with Markov-Switching. *Journal of Econometrics*, **60**, 1–22.

Klaassen, F. (2002) Improving GARCH volatility forecasts with regime-switching GARCH. *Empirical Economics*, **27**, 363–394.

Kubát, M. (2015) Virtual Currency Bitcoin in the Scope of Money Definition and Store of Value. *Procedia Economics and Finance*, **30**, 409–416.

Lamoureux, G. and Lastrapes, W. (1990) Heteroskedasticity in Stock Return Data: Volume versus GARCH Effects. *Journal of Finance*, **45**, 221–229.

Langrock, R. MacDonald, I. L. and Zucchini, W. (2016) *Hidden Markov Models for Time Series*, 2nd ed., Boca Raton, FL, Taylor & Francis Group.

Lee, G and Engle, R. (1999) A permanent and transitory component model of stock return volatility. In *Cointegration Causality and Forecasting A Festschrift in Honor of Clive W. J. Granger*, Oxford, Oxford University Press.

Mandelbrot, B. (1963a) The variation of certain speculative prices. *Journal of Economic Literature*, **30**, 102–146.

Mandelbrot, B. (1963b) New methods in statistical economics. *Journal of Political Economy*, **71**, 421–440.

Mandelbrot, B. (1967) The variation of some other speculative prices. *Journal of Business*, **40**, 393–413.

Marcucci, J. (2005) Forecasting Stock Market Volatility with Regime-Switching GARCH models. *Studies in Nonlinear Dynamics & Econometrics*, **9**(4).

Mehmet, A. (2008) Analysis of Turkish Financial Markets with Markov Regime Switching Volatility Models. The Middle East Technical University.

Mizrach, B. (1990) Learning and Conditional Heteroskedasticity in Asset Returns. *Mimeo*, Department of Finance, The Warthon School, University of Pennsylvania.

Mohd Razali, N. and Yap, B. (2011) Power Comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling Tests. *Journal of Statistical Modeling and Analytics*, **2**(1), 21–33.

Molnár, P. and Thies, S. (2018) Bayesian change point analysis of Bitcoin returns. *Finance Research Letters*, available online 20 March 2018.

Nakamoto, S. (2008) Bitcoin: A Peer-to-Peer Electronic Cash System. Available online at: www.bitcoin.org/bitcoin.pdf. Last accessed: 19th September 2018.

Nelson, D. (1991) Conditional heteroskedasticity in asset returns: A new approach. *Econometrica*, **59**, 347–370.

Nguyen, N. (2017) Hidden Markov Model for Portfolio Management with Mortgage-Backed Securities Exchange-Traded Fund. Committee on Finance Research, Society of Actuaries.

Premanode, B., Sattayatham, P. and Sopipan, N. (2012a) Forecasting Volatility of Gold Price Using Markov Regime Switching and Trading Strategy. *Journal of Mathematical Finance*, **2**, 121–131.

Premanode, B., Sattayatham, P. and Sopipan, N. (2012b) Forecasting Volatility and Price of the SET50 Index using Markov Regime Switching. *Procedia Economics and Finance*, **2**, 265–274.

Reher, G. and Wiling, B. (2011) Markov-switching GARCH models in finance: a unifying framework with an application to the German stock market. Center for Quantitative Economics Working Papers 1711, University of Muenster.

Shaw, C. (2018) Conditional heteroskedasticity in crypto-asset returns. *Journal of Statistics: Advances in Theory & Applications*, (forthcoming).

Sheu, H., Lee, H. and Lai, Y. (2017) A Markov Regime Switching GARCH Model with Realized Measures of Volatility for Optimal Futures Hedging. *Journal of Futures Markets*, **37**(11), 1,124–1,140.

Shi, Y and Ho, K. (2016) A discussion on the Innovation Distribution of Markov Regime-Switching GARCH Model. *Economic Modelling*, **53**, 278–288.

Sims, C. and Zha, T. (2006) Were There Regime Switches in U.S. Monetary Policy? *The American Economic Review*, **96**(1), 54–81.

Smith, R., Sola, M. and Spagnolo, F. (2000) The Prisoner's Dilemma and Regime-Switching in the Greek-Turkish Arms Race. *Journal of Peace Research*, **37**(6), 737–750.

Stock, J. (1998) Eastimating Continuous-Time Processess Subject to Time Deformation. *Journal of the American Statistic Association*, **83**, 77–85.

Sun, P. and Zhou, C. (2014) Diagnosing the distribution of GARCH innovations. *Journal of Empirical Finance*, **29**, 287–303.

Wilfling, B. (2009) Volatility regime-switching in European exchange rates prior to monetary unification. *Journal of International Money and Finance*, **28**, 240–270.