Universidade de São Paulo
Faculdade de Economia, Administração e Contabilidade de
Ribeirão Preto
Departamento de Economia
Programa de Pós-Graduação em Economia — Área: Economia
Aplicada

Pedro Luiz Paolino Chaim

Essays in Financial Econometrics

Ensaaios em Econometria Financeira

Orientador: Márcio Poletti Laurini

Ribeirão Preto

2019

Prof. Dr. Vahan Agopyan
Reitor da Universidade de São Paulo

Prof. Dr. André Lucirton Costa
Diretor da Faculdade de Economia, Administração e Contabilidade de
Ribeirão Preto

Prof. Dr. Sérgio Kannebley Junior
Chefe do Departamento de Economia

Prof. Dr. Sérgio Naruhiko Sakurai
Coordenador do Programa de Pós-Graduação em Economia

PEDRO LUIZ PAOLINO CHAIM

Essays in Financial Econometrics

Ensaaios em Econometria Financeira

Tese apresentada ao Programa de Pós-Graduação em Economia — Área: Economia Aplicada da Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto da Universidade de São Paulo, para obtenção do título de Doutor em Ciências. Versão Corrigida. A original encontra-se disponível na FEA-RP/USP.

Orientador: Márcio Poletti Laurini

Ribeirão Preto

2019

Autorizo a reprodução e divulgação total ou parcial deste trabalho por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001

Chaim, Pedro

Essays in Financial Econometrics/ Universidade de São Paulo – USP

Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto

Departamento de Economia

Programa de Pós-Graduação em Economia - Área: Economia Aplicada; Orientador: Márcio Poletti Laurini

Ribeirão Preto, 2019- 129 p. : il. 30cm

Tese de Doutorado – Universidade de São Paulo, 2019.

1. Bitcoin. 2. Cryptocurrencies. 3. Stochastic Volatility. 4. Co-Jumps. 5. Financial Bubbles. 6. Long Memory. 7. Strict Local Martingales. 8. Filtration Enlargement. 9. Foreign Exchange Markets. 10. Volatility Forecasting. I. Orientador: prof. Dr. Márcio Poletti Laurini. II. Universidade de São Paulo. III. Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto.

Acknowledgements

Agradeço primeiramente à minha família que tornou possível que eu me focasse nos meus estudos. Também aos novos amigos que fiz, que tornaram o curso tão mais agradável.

Ao professor Márcio Poletti Laurini, pela excelente orientação e valiosas dicas. Também agradeço aos professores Fábio Augusto Reis Gomes e Jefferson Donizeti Pereira Bertolai, pelos comentários na minha qualificação e pré-defesa, mas principalmente pelas valiosas lições desde o início do curso.

A todos os professores pelas aulas que tive ao longo do doutorado e que são os responsáveis pelo conhecimento que agreguei ao longo desses processo do doutorado.

Abstract

CHAIM, P. (2019) **Essays in Financial Econometrics**. Doctoral Dissertation - Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto 2019.

This dissertation is composed by five self-contained papers on broad themes in financial econometrics. The first two papers are applications of stochastic volatility models with discontinuous jumps to cryptocurrencies. Third and fourth papers deal with strict local martingale theory, one is an investigation regarding the presence of price bubbles in Bitcoin and the other proposes an interpretation for observed systematic forecast errors in Brazilian foreign exchange markets. The fifth and final paper discusses the estimation of long memory stochastic volatility models using integrated nested Laplace approximations.

Keywords: Bitcoin; Cryptocurrencies; Stochastic Volatility; Co-Jumps; Long Memory; Financial Bubbles; Strict Local Martingales; Filtration Enlargement; Foreign Exchange Markets; Volatility Forecasting.

Resumo

CHAIM, P. (2019) **Ensaaios em Econometria Financeira**. Tese (Doutorado) - Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto 2019.

Esta tese é composta por cinco artigos independentes com temas variados em econometria financeira. Os dois primeiros artigos são aplicações de modelos de volatilidade estocástica com saltos descontínuos para criptomoedas. Os terceiro e quarto artigos lidam com teoria de Martingales locais estritos, um investiga a presença de bolhas no preço do Bitcoin e o outro apresenta uma interpretação para sistemáticos erros de previsão observados no mercado futuro de câmbio brasileiro. O quinto e último artigo discute a estimação de modelos de volatilidade estocástica com memória longa usando aproximações de Laplace integradas aninhadas.

Palavras-chaves: Bitcoin; Criptomoedas; Saltos Conjuntos; Memória Longa; Bolhas Financeiras; Martingales Locais Estritos; Alargamento de Filtração; Mercados Futuros de Câmbio; Predição de Volatilidade.

List of Figures

Figure 2.1 – Bitcoin dollar prices and daily returns	19
Figure 2.2 – Jump process in volatility	22
Figure 2.3 – Jump process in mean	23
Figure 2.4 – Absolute returns and predicted volatility	24
Figure 3.1 – Daily Returns	30
Figure 3.2 – Volatility component and probability of jumps	36
Figure 3.3 – Mean jumps and probability of jumps	37
Figure 3.4 – Unconditional probability of jumps in mean - subsample analysis .	45
Figure 3.5 – Probability of jumps in mean - subsample analysis	46
Figure 3.6 – Common and individual long-run volatility	50
Figure 4.1 – Daily Bitcoin dollar prices and returns	59
Figure 4.2 – Variance function estimates - daily Bitcoin prices	62
Figure 4.3 – Variance function estimates - high-frequency Bitcoin prices	63
Figure 4.4 – Variance function estimates - S&P500, EUR-USD, Gold, and Oil .	64
Figure 4.5 – Fitted Volatility - Andersen and Piterbarg (2007) model	68
Figure 5.1 – Spot rate, Focus and future market forecasts	75
Figure 5.2 – Rolling sample bias test	76
Figure 5.3 – Variance functions, whole sample	78
Figure 5.4 – Vertically rescaled variance functions, whole sample	79
Figure 5.5 – Vertically rescaled variance functions, first subsample	80
Figure 5.6 – Vertically rescaled variance functions, second subsample	80
Figure 5.7 – Posterior Densities, unrestricted model	82
Figure 5.8 – Posterior Densities, restricted model	83
Figure 5.9 – Fitted Volatility	84
Figure 6.1 – Returns - Bitcoin (btc), Ethereum (eth), Litecoin (ltc), S&P 500 (sp500)	101
Figure 6.2 – Spline Component - Bitcoin	106
Figure 6.3 – Bitcoin (btc) - Fitted Volatility and VaR (5%)	107
Figure 6.4 – Ethereum (etc) - Fitted Volatility and VaR (5%)	108
Figure 6.5 – Litecoin (ltc) - Fitted Volatility and VaR (5%)	109
Figure 6.6 – S&P 500 (sp500) - Fitted Volatility and VaR (5%)	110
Figure 6.7 – 10-minutes intraday Bitcoin returns	112
Figure 6.8 – Fitted Intraday Volatility Seasonality	112

Figure 6.9 – Absolute 10-minute Returns and Fitted Intraday Volatility 113

Figure 6.10 –Realized Variance and LMSV-based Realized Variance 113

List of Tables

Table 2.1	– Descriptive statistics of returns	19
Table 2.2	– Standard stochastic volatility parameters posterior distribution . . .	20
Table 2.3	– Parameters posterior distributions Qu and Perron (2013) model . . .	20
Table 2.4	– Parameters posterior distributions Laurini, Mauad and Auibe (2016) model.	21
Table 2.5	– Mean return jumps of largest magnitude	22
Table 3.1	– Cryptocurrency markets	29
Table 3.2	– Descriptive statistics of daily returns	31
Table 3.3	– Correlation between contemporaneous returns	31
Table 3.4	– Descriptive statistics of daily volatility (absolute returns)	32
Table 3.5	– Posterior distributions common parameters	34
Table 3.6	– Posterior distributions individual parameters	35
Table 3.7	– Model Fit Models	38
Table 3.8	– GPH Estimation for the Squared Returns	41
Table 3.9	– GPH Estimation for the Log of Squared Returns	42
Table 4.1	– Descriptive statistics of daily log-returns	58
Table 4.2	– Descriptive statistics of five minutes Bitcoin log-returns	60
Table 4.3	– Posterior Distributions of Andersen and Piterbarg (2007) model - Bitcoin	66
Table 4.4	– Posterior Distributions of Andersen and Piterbarg (2007) model - S&P 500, EUR-USD, Gold, and Oil	67
Table 5.1	– Descriptive statistics	75
Table 5.2	– Martingale tests	76
Table 5.3	– Forecast comparison	77
Table 5.4	– Forecast tests	77
Table 5.5	– Estimation results, unrestricted model	82
Table 5.6	– Estimation results, restricted model	83
Table 5.7	– Model evaluation	84
Table 6.1	– Monte Carlo Experiment LMSV Model - First parameter set ($\mu=-$ 9.9, $H= 0.931$, $\tau=0.962$) - Sample size 500	99
Table 6.2	– Monte Carlo Experiment LMSV Model - First parameter set ($\mu=-$ 9.9, $H= 0.931$, $\tau=0.962$) - Sample size 1000	99

Table 6.3 – Monte Carlo Experiment LMSV Model - Second parameter set ($\mu=-$ 7.182, $H=0.938$, $\tau=0.294$) - Sample size 500	99
Table 6.4 – Monte Carlo Experiment LMSV Model - Second parameter set ($\mu=-$ 7.182, $H=0.938$, $\tau=0.294$) - Sample size 1000	100
Table 6.5 – Descriptive Statistics - Bitcoin (btc), Ethereum (eth), Litecoin (ltc) and S&P 500 (sp500)	100
Table 6.6 – Posterior distribution estimated parameters - Bitcoin (btc)	102
Table 6.7 – Posterior distribution estimated parameters - Ethereum (eth)	103
Table 6.8 – Posterior distribution estimated parameters - Litecoin (ltc)	104
Table 6.9 – Posterior distribution estimated parameters - S&P 500 (sp500)	105
Table 6.10 –Model Fit Measures - Marginal Likelihood and Waic	105
Table 6.11 –In Sample Error Measures	105
Table 6.12 –VaR statistics	106
Table 6.13 –Out of Sample Forecast Measures: Bitcoin and Ethereum	111
Table 6.14 –Out of Sample Forecast Measures: Litecoin and S&P500	114
Table 6.15 –Posterior Estimated Paramaters - LMSV for Bitcoin 10-minute returns	114
Table 6.16 –Forecast Measures for the Realized Variance	115

Contents

1	General Introduction	13
2	Volatility and return jumps in bitcoin	15
2.1	Introduction	15
2.2	Models	17
2.3	Data	18
2.4	Results	18
2.5	Conclusion	22
3	Nonlinear dependence in cryptocurrency markets	25
3.1	Introduction	25
3.2	Data	28
3.3	Model	30
3.4	Results	33
3.4.1	Posterior estimates	34
3.4.2	Simulation Evidence for the Presence of Long Memory Processes	38
3.4.3	Sub-sample Analysis	43
3.5	Conclusion	44
4	Is Bitcoin a bubble?	51
4.1	Introduction	51
4.2	Asset Price Bubbles and Strict Local Martingales	54
4.3	Empirical Evidence	57
4.3.1	Descriptive Statistics	57
4.3.2	Volatility Function	60
4.3.3	Stochastic Volatility Model	61
4.4	Concluding Remarks	68
5	Foreign Exchange Expectation Errors and Filtration Enlargements	69
5.1	Introduction	69
5.2	Filtration Enlargement and Strict Local Martingales	71
5.2.1	Strict Local Martingales	71
5.2.2	Filtration Enlargement	72
5.3	Empirical Evidence	74
5.4	Nonparametric Estimation	78
5.5	Parametric Estimation	79

5.6	Discussion	84
5.7	Concluding Remarks	85
6	Estimating long memory stochastic volatility models using integrated nested Laplace approximations	87
6.1	Introduction	87
6.2	Long Memory Stochastic Volatility Models	91
6.2.1	Long Memory Stochastic Volatility Models	93
6.2.2	Gaussian Markov Random Field approximation of Fractional Gaussian Noise	95
6.2.3	Alternative Specifications	96
6.3	Monte Carlo Simulation Study	98
6.4	Data	99
6.5	Estimation Results - Daily Data	101
6.6	Out-of-sample forecasting	106
6.7	Application to high frequency data	108
6.8	Conclusion	115
6.9	Appendix - Specification Details	116
	Bibliography	119

1 General Introduction

This doctoral dissertation is composed of five self-contained papers in financial econometrics. Two papers are applications of stochastic volatility models with discontinuous jumps to cryptocurrencies. Two papers deal with strict local martingales theory, one investigates the presence of a bubble in Bitcoin, while the other looks at systematic forecast errors in Brazilian foreign markets. These first four papers were published in different journals. The final paper deal with employing INLA (Integrated Nested Laplace Approximations) in estimating stochastic volatility models with long memory and multivariate stochastic volatility models with a common factor structure. It is a shared feature that all papers presented here deal with the Bayesian estimation of stochastic volatility models.

In first paper “Volatility and return jumps in bitcoin” we study the volatility of Bitcoin returns by estimating stochastic volatility models with discontinuous jumps to volatility and to both volatility and mean returns. This paper was published in *Economics Letters* 173 (2018) 158-163.

In the second paper we apply a multivariate version of the stochastic volatility model with jumps to nine cryptocurrencies. This paper was published in the *North American Journal of Economics and Finance* 48 (2019) 32-47, with title “Nonlinear dependence in cryptocurrency markets”.

The third paper investigates the presence of price bubbles in Bitcoin from the point of strict local martingale theory of financial bubbles and is entitled “Is Bitcoin a bubble?”. This paper was published in *Physica-A: Statistical Mechanics and its Applications* 517 (2019) 222-232.

The fourth paper also deals with strict local martingales, but now a filtration enlargement mechanism is pointed as a possible explanation for systematic forecast errors in Brazilian foreign exchange markets. This paper was published in *Stats* 2 (2) 212-227, with the title “Foreign Exchange Errors and Filtration Enlargements”.

The fifth and final paper, still a work in progress, is a discussion on use of INLA to estimate stochastic volatility models with long memory components in volatility. It has the working title “ Estimating long memory stochastic volatility models using integrated nested Laplace approximations”.

2 Volatility and return jumps in bitcoin

Pedro Chaim
Márcio P. Laurini

Abstract

In this article we are interested in understanding the dynamics of Bitcoin daily returns and volatility. Cryptocurrencies have very high unconditional volatility, and are subject to sudden, massive, price swings. We start with a standard log-normal stochastic volatility model, then explore two formulations which incorporate discontinuous jumps to volatility and returns. Jumps to volatility are permanent, while jumps to mean returns have contemporaneous effects only. Results point to two high volatility periods: the first from late 2013 to early 2014, likely associated to the Mt. Gox incident; the second covers the year of 2017, peaking on December — likely driven by increased popular attention. Jumps to mean returns are specially relevant to capture large price variations, mostly negative, associated with formative events in cryptocurrency markets, such as hacks and unsuccessful fork attempts.

Keywords: Bitcoin; Cryptocurrencies; Stochastic Volatility; Jumps.

2.1 Introduction

We are interested in studying the evolution of Bitcoin daily returns and volatility. When contrasting cryptocurrencies returns to more traditional financial assets, two characteristics stand out: very high unconditional volatility, and large occasional price swings. Motivated by these stylized facts we start with a standard log-normal stochastic volatility model, then explore two formulations which incorporate discontinuous jumps to volatility and mean returns.

The model introduced by Qu and Perron (2013), which decomposes the unobserved volatility into a transitory and a permanent component, is of interest here because it allows us to identify structural, long-lasting, changes to cryptocurrency markets. Transitory jumps to returns level incorporated by the formulation of Laurini, Mauad and Auibe (2016) are appropriate to model large abrupt price variations, characteristic of formative events in cryptocurrency markets.

Estimation results suggest introducing a time varying average volatility component reduces temporal persistence in the autoregressive volatility process. The permanent

volatility component highlights two high volatility periods, with a relatively calmer interval in between: from the second half of 2013 to Mt. Gox’s closure in February 2014, and through the notable exposition cryptocurrencies received during 2017, specially by the end of the year. Mean jumps, which affect returns only contemporaneously, account mostly for negative price variations of large magnitude — notably the unsuccessful fork¹ attempt of Bitcoin into Bitcoin XT in August 2015.

There has been increased interest in informational efficiency of cryptocurrency markets recently. Urquhart (2016) evaluates the predictability of Bitcoin’s returns; and, depending on sampling period, finds returns to be either serially uncorrelated or antipersistent. Wei (2018) explores the relation between liquidity and return predictability of 456 cryptocurrencies. He documents returns predictability diminishing as markets grow bigger. Balcilar et al. (2017) try to predict Bitcoin returns and volatility from trading volume data. They argue transaction volume can sometimes help predict returns, but convey no information on volatility. Katsiampa (2017) fits several GARCH models to Bitcoin volatility and finds including both a short-run and a long-run component of conditional variance is important. Bouri et al. (2017a) and Bouri et al. (2017b) show Bitcoin can be useful for hedging against commodity indexes and uncertainty measures.

The impressive surge in cryptocurrency prices in recent years fuelled the bubble narrative. Cheah and Fry (2015) analyze daily Bitcoin prices from 2011 to 2014. They argue the sharp rise in late 2013 followed by a bust was a bubble. MacDonell (2014) fits a log-periodic law to Bitcoin prices, and argues it ex-ante predicts the crash in December 2013. Corbet, Lucey and Yarovaya (2018) relate Bitcoin and Ethereum prices to “fundamental drivers” to date several bubble periods, including late 2013 and the second semester of 2017.

There is some empirical evidence Bitcoin returns volatility displays long-memory characteristics (Tiwari et al. (2018), Bariviera (2017)). It is understood that the presence of level shifts in a time series might appear as increased persistence (Diebold and Inoue (2001)). The volatility decomposition into transitory and permanent of Qu and Perron (2013) is well suited to deal with such complication.

Discontinuous jumps have been recognized as important features to describe the dynamics of cryptocurrencies. In an early study, Gronwald (2014) applies an autoregressive jump-intensity GARCH model to Bitcoin prices. Scaillet, Treccani and Trevisan (2017) take advantage of the database leak of Mt. Gox to analyze the dynamics of Bitcoin prices in a high-frequency framework. Their results suggest jumps are frequent and clustered in time.

¹ In very general terms, a “fork” is a change to a cryptocurrency internal protocol, which might, or might not, invalidate the record of past transactions. (see <https://www.investopedia.com/terms/h/hard-fork.asp>, <https://www.investopedia.com/terms/h/soft-fork.asp>).

This paper is structured as follows. Section 2.2 presents the models to be estimated, Section 2.3 describes the data, Section 2.4 presents and discusses our main results, Section 2.5 concludes.

2.2 Models

A standard log-normal stochastic volatility model is often specified by the pair of equations

$$y_t = \exp\left(\frac{h_t}{2}\right) \varepsilon_t, \quad \varepsilon_t \sim N(0, 1), \quad (2.1)$$

$$h_t = \mu + \phi(h_{t-1} - \mu) + \sigma^h z_t^h, \quad z_t^h \sim N(0, 1). \quad (2.2)$$

Here y_t are daily log-returns, ε_t is a normal disturbance which determines returns sign and magnitude; h_t is the unobserved log-volatility, which follows a first-order autoregressive process with mean μ , persistence ϕ and shock standard deviation σ^h .

Qu and Perron (2013) decompose log-volatility into a zero-mean transitory component h_t , and a permanent component μ_t . The permanent volatility component μ_t is subject to a normal disturbance z_t^v , with standard deviation σ^v , if the Bernoulli process δ_t^v is a success.

$$y_t = \exp\left(\frac{\mu_t}{2} + \frac{h_t}{2}\right) \varepsilon_t, \quad \varepsilon_t \sim N(0, 1), \quad (2.3)$$

$$h_t = \phi h_{t-1} + \sigma^h z_t^h, \quad z_t^h \sim N(0, 1), \quad (2.4)$$

$$\mu_t = \mu_{t-1} + \delta_t^v \sigma^v z_t^v, \quad z_t^v \sim N(0, 1), \quad (2.5)$$

$$\delta_t^v \sim \text{Bernoulli}(p^v). \quad (2.6)$$

This decomposition of unobserved variance can be viewed as a regime shift model, in which both the number of regimes and the probability of a shift are endogenously determined.

Laurini, Mauad and Auibe (2016) further explore the possibility of jumps, introducing a jump to the mean of the return process y_t , which affects it only contemporaneously. The key idea is that if the realization of the Bernoulli random variable δ_t^m is a success, then the normal random variable ε_t is drawn from $N(\gamma_t, 1)$, instead of $N(0, 1)$, where γ_t is the size of the mean jump (which depends on a normal disturbance z_t^m with

standard deviation σ^m). A univariate version of their model can be written as

$$y_t = \exp\left(\frac{\mu_t}{2} + \frac{h_t}{2}\right) \varepsilon_t, \quad \varepsilon_t \sim N(\gamma_t, 1), \quad (2.7)$$

$$h_t = \phi h_{t-1} + \sigma^h z_t^h, \quad z_t^h \sim N(0, 1), \quad (2.8)$$

$$\mu_t = \mu_{t-1} + \delta_t^v \sigma^v z_t^v, \quad z_t^v \sim N(0, 1), \quad (2.9)$$

$$\gamma_t = \delta_t^m \sigma^m z_t^m, \quad z_t^m \sim N(0, 1), \quad (2.10)$$

$$\delta_t^v \sim \text{Bernoulli}(p^v), \quad \delta_t^m \sim \text{Bernoulli}(p^m). \quad (2.11)$$

This mean jump feature is appropriate to capture abrupt price swings, characteristic of highly speculative assets such as cryptocurrencies.

All models are estimated using Bayesian methods, using Markov Chain Monte Carlo procedures, through a composition of Gibbs, Metropolis-Hastings and a threshold sampling scheme using an auxiliary variable to estimate the jumping processes, as described in Laurini, Mauad and Auibe (2016). We use the same structure of priors in that article. Due to space issues we do not present the estimation procedure and the structure of priors, and so we indicate the reading of these details in Laurini, Mauad and Auibe (2016).

2.3 Data

Table 2.1 displays descriptive statistics of Bitcoin returns. In order to provide context, we also show these statistics for the returns of gold, the USD-EUR exchange rate, and the S&P500 index. Sample covers April 2013 through May 2018, totaling 1840 daily return observations. We chose to leave out earlier periods due to low market liquidity².

It is immediately evident that Bitcoin is much more volatile than traditional financial assets, with average daily price variation of 4.40% — compare it to S&P500's 0.66%. Interestingly, higher moments are not so extreme. For our sample period of May 2016 to May 2018, Bitcoin's returns are rather symmetrical, and have a smaller kurtosis coefficient than S&P500.

2.4 Results

Here we describe and discuss the implications of our main results. First we look at parameter estimates, then the evolution of jumps to volatility and returns, and finally their impact on predicted volatility.

² Bitcoin dollar price data sourced from <http://coinmetrics.io/data-downloads>. Data on the other assets from <http://fred.stlouisfed.org>.

Table 2.1 – Descriptive statistics of returns

	Mean	Std.	Min	q05	q50	q95	Max	Skew.	Kurt.
Bitcoin	0.396105	4.4088	-20.208	-7.0550	0.37268	6.9655	22.351	-0.10220	6.789
Gold	0.001028	0.7438	-2.722	-1.1733	0.01444	1.2182	3.730	0.07085	5.046
EUR-USD	0.007105	0.4807	-2.672	-0.6796	-0.01626	0.7786	1.559	-0.35661	5.402
S&P500	0.053743	0.6607	-4.716	-0.9220	0.09091	0.9021	2.871	-1.13176	12.401

Note: table reports descriptive statistics of daily log-returns Bitcoin, Gold, the EUR-USD exchange rate, and the S&P500 index. First column display the mean. Second column shows the standard deviation. Third and seventh columns show the smallest and largest observations, respectively. Fourth through sixth show 0.5%, 50%, and 95% quantiles, respectively. The seventh column reports the skewness coefficient. Last column reports the (raw) kurtosis coefficient.

Figure 2.1 – Bitcoin dollar prices and daily returns

Note: figure displays Bitcoin dollar prices in top panel (a), and daily returns in bottom panel (b). Price data was collected from <https://coinmetrics.io/data-downloads/>.

Table 2.2 summarizes posterior parameter estimates of the standard stochastic volatility model (2.1)-(2.2). Long run log-volatility μ has posterior mean -7.16 , which implies $\exp(-7.16/2) = 2.78\%$ daily price variation. Autoregressive parameter ϕ is estimated around 0.90 and 0.95 with 90% probability — which suggests substantial temporal persistence of h_t . The standard deviation of volatility, σ^h is estimated between 0.45 and 0.60 with 90% probability.

Table 2.3 displays descriptive statistics of the posterior distributions of Qu and Perron (2013) model (2.3)-(2.6) parameters. As expected, the introduction of time varying average volatility μ_t reduces temporal persistence of h_t . The autoregressive parameter ϕ is now estimated between 0.65 and 0.90 with 90% probability. We estimate the unconditional

Table 2.2 – Standard stochastic volatility parameters posterior distribution

	Mean	Std.	q05	q50	q95	Skew.	Kurt.
μ	-7.1628	0.1833	-7.4659	-7.1596	-6.8719	-0.2132	3.7189
ϕ	0.9296	0.0141	0.9052	0.9303	0.9517	-0.2892	3.1304
σ^h	0.5242	0.0456	0.4515	0.5231	0.6015	0.1559	3.0242

Note: table presents descriptive statistics and posterior distributions of the standard stochastic volatility model (2.1)-(2.2) parameters. First column displays posterior means. Second column displays posterior standard deviations. Third column shows 5% percentiles. Fourth column shows the medians.

Fifth column shows 95% posterior percentiles. Sixth column shows skewness coefficients. Seventh column shows raw kurtosis coefficients. Posterior sampling was carried out with 20000 MCMC iterations using the stochvol R package of Kastner (2016).

Table 2.3 – Parameters posterior distributions Qu and Perron (2013) model

	Mean	Std.	q05	q50	q95	Skew.	Kurt.
ϕ	0.8018	0.0661	0.6551	0.8188	0.8762	-1.2319	4.1908
σ^h	0.7119	0.0848	0.5954	0.7032	0.8799	0.7242	3.3828
σ^v	0.5406	0.1593	0.348	0.5138	0.8439	0.9228	3.4535
p_v	0.0467	0.0278	0.0136	0.0389	0.1	0.7675	2.2859

Note: table presents descriptive statistics and selected quantiles of posterior distributions of model (2.7)-(2.11) parameters (rows). First column displays posterior means. Second column displays posterior standard deviations. Third column shows 5% percentiles. Fourth column shows the medians. Fifth column shows 95% posterior percentiles. Sixth column shows skewness coefficients. Seventh column shows raw kurtosis coefficients. Posterior sampling was carried out with 20000 MCMC iterations, through the method described by Laurini, Mauad and Auibe (2016).

probability of a permanent jump to volatility p^v with posterior mean of 0.0467.

Table 2.4 presents the posterior distributions of the parameters of the Laurini, Mauad and Auibe (2016) double jump stochastic volatility model (2.7)-(2.11). Persistence ϕ is also lower than in the standard stochastic volatility case. Unconditional probability of a jump to volatility p^v is smaller, with posterior mean 0.0364; likely because some variation is explained by transitory mean jumps γ_t — which has mean posterior unconditional probability p^m of 0.0142. We will see below that mean returns jumps account mainly for large negative price variations.

Figure 2.2 displays the evolution of the permanent volatility component μ_t in panel (a), and the conditional probability of a jump to volatility in panel (b). The dashed horizontal line in panel (a) represents the long-run average log-volatility μ , as implied by the standard stochastic volatility model (2.1)-(2.2) — whose parameter estimates are provided in Table 2.2³. The permanent volatility component μ_t was higher than constant long-run average log-volatility μ from the start of our sample to May 2014, then decreased from early 2015 to a minimum in October 2016. It then increased steeply from early 2017, to a peak in December 2017.

³ We only show μ_t implied by the double jump model because both estimates are very similar.

Table 2.4 – Parameters posterior distributions Laurini, Mauad and Auibe (2016) model.

	Mean	Std.	q05	q50	q95	Skew.	Kurt.
ϕ	0.8383	0.042	0.7666	0.8395	0.8984	-1.1885	8.7544
σ^h	0.6405	0.0909	0.4786	0.6454	0.7844	0.0213	3.3173
σ^v	0.5523	0.1706	0.356	0.493	0.906	0.8692	2.9161
σ^m	0.7274	0.2855	0.3099	0.7544	1.2205	0.2634	2.1259
p_v	0.0364	0.0269	0.0086	0.026	0.0956	1.191	3.2118
p_m	0.0142	0.0259	4.00E-04	0.0024	0.0934	2.3226	7.3285

Note: table presents descriptive statistics and selected quantiles of posterior distributions of model (2.7)-(2.11) parameters (rows). First column displays posterior means. Second column displays posterior standard deviations. Third column shows 5% percentiles. Fourth column shows the medians. Fifth column shows 95% posterior percentiles. Sixth column shows skewness coefficients. Seventh column shows raw kurtosis coefficients. Posterior sampling was carried out with 20000 MCMC iterations, through the method described by Laurini, Mauad and Auibe (2016).

A first high volatility period took place from late 2013 to early 2014, and was most likely associated to the Mt. Gox incident⁴. Bitcoin dollar prices peaked at about \$1,100, on November 29, 2013, then fell below \$500 by April 2014. Mt. Gox was the largest Bitcoin exchange, handling approximately 70% of all transactions. During this period, \$450 million worth of Bitcoins mysteriously vanished from users digital wallets. On February 7, 2014, Mt. Gox halted all Bitcoin withdrawals. There is evidence that fraudulent inside activity drove Bitcoin's price surge between February and November 2013 (Gandal et al. (2018)). This seems to be a very relevant episode in the history of Bitcoin.

In the aftermath of Mt. Gox's closure in February 2014, volatility shortly spiked, then fell below the unconditional long-run average (μ , the dashed line in Figure 2.2, panel (a)). Bitcoin prices were hit hard by the incident, and it was not until January 2017 that the Bitcoin-USD exchange rate again exceeded the \$1,000 mark.

The year of 2017 was a period of increased popular interest in cryptocurrencies. From January 1st, 2017, to January 1st, 2018, Bitcoin dollar prices went from \$963 to \$14,112, an average of 0.7% daily price increase. Through 2017, the permanent volatility component increased steadily — by the end of the year, reached the same high levels of 2013.

Figure 2.3 displays mean jumps γ_t in panel (a), and the probability of a mean jump in panel (b). These transitory jumps to returns mean seem to be specially relevant to capture large negative price variations since all γ_t with magnitude superior to 0.05 are negative — larger mean jumps are dated in Table 2.5. The return jump γ_t of largest magnitude is negative and occurred on August 19, 2015 (not counting the large jump in the

⁴ The whole affair was dragged from February 2013 to February 2014, see <https://www.wired.com/2014/03/bitcoin-exchange/> for an account.

Table 2.5 – Mean return jumps of largest magnitude

Dates	Price	y_t	μ_t	γ_t
2013-05-02	116.38	-0.1776132	-6.874563	-0.17336225
2014-03-28	477.14	-0.1956663	-6.942225	-0.07878721
2015-08-19	225.67	-0.1336145	-7.856671	-0.12824021
2016-01-16	365.07	-0.1642773	-7.747274	-0.07487134
2017-01-12	775.18	-0.1582813	-7.624183	-0.06896139

Note: table shows Bitcoin price, returns y_t , average volatility μ_t , and mean return jumps γ_t , at the dates in which mean return jumps of largest magnitude happened.

Figure 2.2 – Jump process in volatility

Note: figure displays the evolution of the permanent volatility component μ_t in panel (a), and the probability of a jump to volatility in panel (b). We only show μ_t implied by the double jump model because both estimates are very similar.

first period). This is likely associated to the introduction of Bitcoin XT, an unsuccessful attempt to fork of the original digital currency that started on August 15, 2015.

Figure 2.4 plots Bitcoin absolute returns (gray), predicted volatility (orange), and average volatility $\exp(\mu_t/2)$ (black). We can see in terms of absolute returns, how much average log-volatility changes from the two high volatility episodes of 2013 and 2017, when compared to the relatively calmer period between mid 2015 and early 2017.

2.5 Conclusion

We have explored the evolution of Bitcoin daily returns and volatility from May 2013 to April 2018. This asset displays patterns of varying average volatility and discon-

Figure 2.3 – Jump process in mean

Note: figure displays transitory mean jumps in panel (a), and the occurrence probability of a mean jump in panel (b).

tinuous return jumps that are not properly captured by traditional models of conditional volatility. Results indicate the importance of incorporating permanent jumps to volatility, such as in the structure proposed by Qu and Perron (2013), and also the relevance of transitory jumps in mean returns, as proposed by Laurini, Mauad and Auibe (2016).

Parameter estimates show that introducing a time varying average volatility component reduces temporal persistence of the autoregressive volatility process. The permanent volatility component points to two high volatility periods, with a relatively calmer interval in between: from the second half of 2013 to Mt. Gox's closure in February 2014, and through the notable exposition cryptocurrencies received during the year of 2017.

Mean jumps, which affect returns only contemporaneously, account mostly for negative price variations of large magnitude — notably the unsuccessful fork attempt of Bitcoin into Bitcoin XT in August 2015.

These results are potentially relevant for portfolio and risk management procedures, since the presence of these jumps can substantially impact the structure of losses and gains related to this asset.

Figure 2.4 – Absolute returns and predicted volatility

Note: figure plots Bitcoin absolute returns (gray), predicted volatility (orange) (sum of permanent and transitory volatility components), and permanent volatility component $\exp(\mu_t/2)$ (black) estimated using the double jump model.

3 Nonlinear dependence in cryptocurrency markets

Pedro Chaim
Márcio P. Laurini

Abstract

We are interested in describing the returns and volatility dynamics of major cryptocurrencies. Very high volatility, large abrupt price swings, and apparent long memory in volatility are documented features of such assets. We estimate a multivariate stochastic volatility model with discontinuous jumps to mean returns and volatility. This formulation allows us to extract a time-varying shared average volatility and to account for possible large outliers. Nine cryptocurrencies with roughly three years of daily price observations are considered in the sample. Our results point to two high volatility periods in 2017 and early 2018. Qualitatively, the permanent volatility component seems driven by major market developments, as well as the level of popular interest in cryptocurrencies. Transitory mean jumps become larger and more frequent starting from early 2017, further suggesting shifts in cryptocurrencies return dynamics. Calibrated simulation exercises suggest the long memory dependence features of cryptocurrencies are well reproduced by stationary models with jump components.

Keywords: Bitcoin; Cryptocurrencies; Risk; Volatility; Co-Jumps; Long Memory.

3.1 Introduction

We are interested in exploring the joint dynamics of major cryptocurrencies. Recent attention devoted to statistical analysis of cryptocurrency markets has identified preeminent features, such as very high volatility, large return outliers, and increased temporal dependency of volatility. Another important aspect is the presence of long memory dependence structures in this asset class. In this work we use a multivariate stochastic volatility model with jumps in the mean and variance that reproduces the empirical patterns observed in the mean and conditional volatility of these series and is also able to generate the long memory behavior, reproducing all the essential aspects observed in this market.

Several recent empirical studies focus on the dynamics of volatility of cryptocurrencies returns. Generalized Autoregressive Conditional Heteroskedasticity (GARCH)

models are a popular approach. Dyhrberg (2016b), Dyhrberg (2016a) estimates asymmetric threshold GARCH models to Bitcoin and argues the asset has hedging properties against stocks in the Financial Times Stock Exchange. Katsiampa (2017) compares several GARCH specifications to Bitcoin and finds an AR-CGARCH, which incorporates a long-run volatility component, to give best fit. Robustness of the results presented by Katsiampa (2017) and Dyhrberg (2016a) are questioned by Charles and Darné (2019) and Baur, Dimpfl and Kuck (2018), respectively. Baur and Dimpfl (2018) consider Bitcoin and five other cryptocurrencies in their analysis, an integrated GARCH model is shown to have best fit (in terms of likelihood criteria) for four cryptocurrencies; they also perform unconditional and conditional Value at Risk coverage tests. Return and volatility spillovers among cryptocurrencies are studied in Koutmos (2018) and Katsiampa (2018), and tail risk measures in Gkillas and Katsiampa (2018) and Troster et al. (2018). The presence of possible jumps and regime changes was studied by Chaim and Laurini (2018), Fry (2018) and Ardia, Bluteau and Rüede (2018).

Long range dependence in cryptocurrency volatility is a well-documented feature. Bariviera et al. (2017) calculate Hurst exponents for Bitcoin returns volatility and argue there is evidence of self-similarity. Lahmiri, Bekiros and Salvi (2018) provide evidence of long-range dependence in seven Bitcoin markets using a fractionally integrated GARCH model under several specifications of error distribution. As discussed by Diebold and Inoue (2001) and Lai and Xing (2006), long range dependence characteristics can appear due to structural breaks affecting parameters of processes driving conditional volatility and returns. Charfeddine and Maouchi (2018) employ several tests to argue cryptocurrency volatility has indeed true long memory rather than level shifts. The relationship between long memory and market inefficiency for cryptocurrencies was studied in Urquhart (2016) and Cheah et al. (2018). Other aspects related to the presence of long memory in this market can be found in Phillip, Chan and Peiris (2018), Al-Yahyaee, Mensi and Yoon (2018), Jiang, Nie and Ruan (2018), Alvarez-Ramirez, Rodriguez and Ibarra-Valdez (2018) and Zargar and Kumar (2019).

In this work we show that the presence of level shifts in the structure of volatility, generated by jump processes, can reproduce the long memory behavior observed in these assets. A key aspect of data generating processes employed in this work is that series are constructed by mixing short memory processes with time varying conditional volatility, as opposed to true long memory processes. This is relevant in terms of volatility predictability, portfolio allocation, and risk management, since different underlying persistence properties are implied.

Our sample consists of nine major cryptocurrencies with roughly three years of daily price observations, ranging from August 2015 to October 2018. Descriptive statistics in Section 3.2 illustrate cryptocurrency characteristics such as an overall very high

level of unconditional volatility, the presence of large outliers, and mostly right-skewed, leptokurtic distributions of returns. Contemporaneous correlation between daily returns are positive and volatilities display signs of long-range temporal dependence, reproducing all the stylized empirical facts observed for these assets.

Since optimal allocation applications and risk measures depend on model choices, the selected specification should account for peculiarities of cryptocurrencies. The multivariate stochastic volatility model introduced by Laurini, Mauad and Auibe (2016) is suitable to deal with those characteristics. Conditional volatility is decomposed in an individual, mean-reverting, component; and a common permanent component which varies discontinuously when subject to a jump. Additionally, the level of returns is also subject to common jumps, but with contemporaneous effect only. This model can be viewed as a regime shift model in which both the number of regimes and transition probabilities are endogenously determined.

Due to the presence of several latent variables model estimation is performed through a mixed Markov Chain Monte Carlo sampling procedure. The results obtained by the estimation of the model indicate that common time varying average daily volatility remains stable around 1.83% from 2015 until March 2017, when it jumps to a high level of about 3.87% daily volatility. This high volatility episode lasts until June. From June to November, average volatility rescinds to about 2.73%, still above pre-2017 levels. From late November to January 2018, amid great popular interest, volatility in cryptocurrency markets again increases. There is a steep reduction in volatility towards the end of our sample. This implied path of volatility is similar to what one finds if the model is estimated for each cryptocurrency individually. Unconditional probability of transitory mean return jumps is estimated 0.35, and those jumps become larger and more frequent starting from 2017.

Employing estimated posterior means as calibration values for data generating processes, we present some Monte Carlo experiments showing that the multivariate stochastic volatility models with jumps in the mean and variance can generate the long memory patterns observed in these assets. The results obtained in our analysis indicate the importance of the presence of jumps and changes of level as components of conditional mean and volatility in econometric modelling of cryptocurrency market.

The remainder of this paper is structured as follows. Section 3.2 presents the data and selected descriptive statistics. Section 3.3 describes the model and estimation procedure. Section 3.4 presents and discusses the main results. Section 3.5 concludes.

3.2 Data

Table 3.1 ranks cryptocurrencies in descending order of market capitalization. As of May 6, 2015, the entire cryptocurrency market consisted of about 400\$ billion. Bitcoin concentrates 36% of the market, and Ethereum is second in market share, with 17%. Cryptocurrencies with over 1\$ billion market capitalization, displayed in Table 3.1, accounted for about 88% of the whole market. Data availability on Ethereum limits our sample. We consider cryptocurrencies which surpassed 1\$ billion in market capitalization over the cryptocurrency boom of 2017, and have roughly three years of daily price observations. These criteria leave us with nine cryptocurrencies: Bitcoin, Ethereum, Ripple, Litecoin, Stellar, Dash, Monero, NEM, and Verge¹; with 1174 daily price observations between August 16, 2015 to October 31, 2018.

Descriptive statistics of daily log returns are displayed in Table 3.2. Returns have positive mean over the considered sample, which mostly cover periods of cryptocurrency expansion. Standard deviations are high relative to more traditional financial assets, as is characteristic of this particular asset class. Verge is the most volatile cryptocurrency, with daily returns varying, on average, 16.5%; and Bitcoin, the most consolidated digital currency, has daily returns varying, on average, 4%. Except for Bitcoin, cryptocurrencies display right skewed returns, with mean larger than median. Daily returns are leptokurtic, as is expected from financial assets, and some cryptocurrencies have very large outlier realizations, which increase kurtosis coefficients. Figure 3.1 plots daily log returns of the nine cryptocurrencies in our sample, it is clear some of the assets, such as Ripple, Litecoin, Monero, and NEM, present very large outliers — possibly due to formative events in those particular markets.

Table 3.3 shows correlation between contemporaneous daily returns of digital currencies in our sample. Coefficients are all positive; correlation is overall highest with Bitcoin, and smallest with Verge. This is some preliminary evidence of cryptocurrency returns displaying contemporaneous comovements.

Table 3.4 presents selected descriptive statistics and dependence measures of returns volatility, measured by the absolute returns. Column 7 reports p-values of Augmented Dickey Fuller (ADF) unit root tests, and column 8 reports p-values of Kwiatkowski et al. (1992) (KPSS) stationarity tests. ADF’s null hypotheses of non-stationarity are uniformly rejected at 1% confidence level. Conversely, KPSS’s null hypotheses of stationarity are also all rejected at 1%. Estimates of the fractional difference parameter d are between 0.29 and 0.40, which falls in the stationary range, but far from what we would expect from

¹ Data was collected from (coinmetrics.io/data-downloads)(access on October 31, 2018). Ethereum limits our sample size, with earliest daily price observation dating August 07, 2015. We exclude the first eight observations due the destabilizing impact of extreme Ethereum return realizations in those days have on empirical estimates. Daily closing prices are computed at 00:00.

Table 3.1 – Cryptocurrency markets

Rank	Name	Symbol	Market Cap.	% Mkt. Cap.	Price
1	Bitcoin	BTC	\$165,947,375,277	36.07%	\$9,751.33
2	Ethereum	ETH	\$79,954,945,093	17.38%	\$805.52
3	Ripple	XRP	\$34,964,625,007	7.60%	\$0.89
4	Bitcoin Cash	BCH	\$30,167,524,727	6.56%	\$1,762.90
5	EOS	EOS	\$14,461,111,605	3.14%	\$17.26
6	Litecoin	LTC	\$9,893,568,748	2.15%	\$175.41
7	Cardano	ADA	\$9,180,172,357	2.00%	\$0.35
8	Stellar	XLM	\$7,819,698,649	1.70%	\$0.42
9	IOTA	MIOTA	\$6,597,460,680	1.43%	\$2.37
10	TRON	TRX	\$5,531,270,046	1.20%	\$0.08
11	NEO	NEO	\$5,510,690,419	1.20%	\$84.78
12	Dash	DASH	\$4,030,203,916	0.88%	\$500.33
13	Monero	XMR	\$3,824,282,820	0.83%	\$238.98
14	NEM	XEM	\$3,747,774,940	0.81%	\$0.42
15	Vechain	VEN	\$2,646,691,078	0.58%	\$5.03
16	Ethereum Classic	ETC	\$2,464,362,518	0.54%	\$24.27
17	Tether	USDT	\$2,260,646,694	0.49%	\$1.00
18	Qtum	QTUM	\$2,011,712,156	0.44%	\$22.71
19	OmiseGO	OMG	\$1,750,481,140	0.38%	\$17.15
20	ICON	ICX	\$1,653,755,856	0.36%	\$4.27
21	Binance Coin	BNB	\$1,647,292,567	0.36%	\$14.44
22	Lisk	LSK	\$1,391,739,656	0.30%	\$13.18
23	Bitcoin Gold	BTG	\$1,359,545,417	0.30%	\$80.04
24	Bytecoin	BCN	\$1,325,028,027	0.29%	\$0.01
25	Zcash	ZEC	\$1,164,652,238	0.25%	\$302.14
26	Nano	NANO	\$1,148,466,740	0.25%	\$8.62
27	Verge	XVG	\$1,138,578,906	0.25%	\$0.08
28	Zilliqa	ZIL	\$1,039,166,283	0.23%	\$0.14
29	Aeternity	AE	\$1,021,907,254	0.22%	\$4.39
30	Ontology	ONT	\$1,000,404,503	0.22%	\$8.88
Total			\$460,085,546,696		

Note: Table ranks cryptocurrencies in terms of market capitalization, as of 2018-05-06. Assets included in our sample are highlighted in bold. Market capitalization data collected from <https://coinmarketcap.com/historical/20180506/>.

Figure 3.1 – Daily Returns

Note: figure plots daily log returns of the nine cryptocurrencies in our sample, from 2015-08-16 to 2018-10-31.

short memory processes. Contrasting test indications regarding stationarity, as well as estimates of fractional difference parameters, reinforce the established notion that returns volatility of cryptocurrencies display long range dependence characteristics. As discussed by Diebold and Inoue (2001) and Lai and Xing (2006), long range dependence characteristics can appear due to structural breaks affecting parameters driving returns and volatility dynamics.

3.3 Model

In this section we describe the multivariate stochastic volatility proposed by Laurini, Mauad and Auibe (2016), which features common jumps to the mean and volatility of the return process. The model is given by the following hierarchical representation:

Table 3.2 – Descriptive statistics of daily returns

	Mean	Std.	Min	2.5q	50q	97.5q	Max	Skew.	Kurt.
Bitcoin	0.271	3.921	-20.208	-8.586	0.279	8.578	22.351	-0.154	7.832
Ethereum	0.400	6.618	-31.984	-12.826	-0.090	15.538	28.629	0.184	6.135
Ripple	0.341	7.688	-60.171	-11.820	-0.311	17.242	101.096	2.974	39.656
Litecoin	0.217	5.757	-39.105	-10.650	0.000	11.817	51.845	1.327	16.474
Stellar	0.392	8.441	-33.342	-14.453	-0.301	18.855	70.404	2.100	18.105
Dash	0.333	5.936	-24.343	-11.499	-0.121	14.101	38.310	0.840	8.170
Monero	0.442	7.138	-29.173	-13.223	0.000	15.623	56.767	1.012	10.331
NEM	0.562	9.203	-43.083	-14.276	0.000	20.152	106.849	2.206	23.276
Verge	0.565	16.538	-69.315	-31.366	0.000	38.621	95.572	0.641	7.649

Note: table presents descriptive statistics of daily log returns of the nine cryptocurrencies in our sample (which goes from 2015-08-16 to 2018-10-31). We report returns in percent points for better presentation. First column displays the mean. Second column shows standard deviations. Third and seventh columns show smallest and largest observations, respectively. Fourth through sixth columns show 2.5%, 50%, and 97.5% quantiles, respectively. Seventh column reports skewness coefficients, and the eighth column presents (raw) kurtosis coefficients.

Table 3.3 – Correlation between contemporaneous returns

	Bitcoin	Ethereum	Ripple	Litecoin	Stellar	Dash	Monero	NEM	Verge
Bitcoin	1.000	0.376	0.277	0.591	0.340	0.453	0.465	0.366	0.238
Ethereum	0.376	1.000	0.250	0.371	0.269	0.386	0.376	0.274	0.155
Ripple	0.277	0.250	1.000	0.336	0.539	0.220	0.277	0.280	0.115
Litecoin	0.591	0.371	0.336	1.000	0.355	0.419	0.406	0.369	0.147
Stellar	0.340	0.269	0.539	0.355	1.000	0.276	0.369	0.370	0.158
Dash	0.453	0.386	0.220	0.419	0.276	1.000	0.463	0.320	0.210
Monero	0.465	0.376	0.277	0.406	0.369	0.463	1.000	0.288	0.186
NEM	0.366	0.274	0.280	0.369	0.370	0.320	0.288	1.000	0.206
Verge	0.238	0.155	0.115	0.147	0.158	0.210	0.186	0.206	1.000

Note: table reports Pearson correlation coefficients between contemporaneous daily log-returns of cryptocurrencies in our sample.

$$y_{i,t} = \exp\left(\frac{h_{i,t}}{2} + \frac{s_i^v \mu_t}{2}\right) \varepsilon_{i,t}, \quad \varepsilon_{i,t} \sim N(\gamma_{i,t}, 1) \quad (3.1)$$

$$h_{i,t} = \phi_i h_{i,t-1} + \sigma_i^h z_{i,t}^h, \quad z_{i,t}^h \sim N(0, 1), \quad (3.2)$$

$$\mu_t = \mu_{t-1} + \delta_t^v \sigma^v z_t^v, \quad z_t^v \sim N(0, 1), \quad (3.3)$$

$$\gamma_{i,t} = s_i^m \delta_t^m \sigma^m z_t^m, \quad z_t^m \sim N(0, 1), \quad (3.4)$$

$$\delta_t^v \sim \text{Bernoulli}(p_v), \quad \delta_t^m \sim \text{Bernoulli}(p_m), \quad (3.5)$$

$$h_{(i,j),t} = s_i^v s_j^v \mu_t, \quad (3.6)$$

where $y_{i,t}$ is the log return of the i -th cryptocurrency in time t . The conditional variance in equation (3.1) is decomposed into an idiosyncratic transitory component $h_{i,t}$, which follows a zero-mean first-order autoregressive process with persistence ϕ_i , and error standard

Table 3.4 – Descriptive statistics of daily volatility (absolute returns)

	Mean	Std.	2.5q	50q	97.5q	d-GPH	ADF	KPSS
Bitcoin	2.538	2.985	0.057	1.382	10.865	0.395	0.010	0.010
Ethereum	4.542	4.821	0.116	2.945	18.387	0.330	0.010	0.010
Ripple	4.123	6.491	0.103	2.103	19.893	0.321	0.010	0.010
Litecoin	3.458	4.609	0.080	1.812	15.460	0.392	0.010	0.010
Stellar	5.189	6.665	0.142	3.205	22.018	0.296	0.010	0.010
Dash	4.085	4.314	0.112	2.839	15.730	0.366	0.010	0.010
Monero	4.872	5.213	0.146	3.273	17.900	0.357	0.010	0.010
NEM	5.867	7.102	0.214	3.863	22.317	0.206	0.010	0.010
Verge	10.827	12.510	0.221	6.846	43.559	0.297	0.010	0.010

Note: table displays selected descriptive statistics of daily log returns volatility, as measured by the absolute returns, of the nine cryptocurrencies in our sample (columns one through five), Geweke and Porter-Hudak (1983) long memory parameter d estimates (sixth column), p-values of unit root ADF tests (seventh column), and p-values of stationarity KPSS tests (eight column); p-values of 0.01 mean lower than 0.01. Volatility is reported in percent points for better presentation.

deviation σ_i^h ; and a permanent component μ_t , scaled by a factor loading s_i^v . The μ_t component can be interpreted as a common (global) variance factor subjected to jumps modelled by a compound binomial process, similar to the one employed by Qu and Perron (2013). Additionally, returns $y_{i,t}$ are subjected to a temporary jump in mean, $\gamma_{i,t}$ — which is incorporated through the mean of the error process $\varepsilon_{i,t}$. Note that volatility jumps have permanent effects, while mean jumps only affect the return process contemporaneously. Incorporating jumps to the return process is a way to considering large outliers.

The permanent component μ_t is subjected to a Gaussian iid disturbance with zero mean and standard deviation σ^v if the realization of the Bernoulli random variable δ_t^v (with parameter p_v) is one. The Bernoulli process captures the occurrence of jumps in common volatility. If this process realization is zero, the process maintains the previous value, and thus the common volatility level is constant. If a jump occurs, the change is given by the Gaussian innovation with standard deviation σ^v . This formulation allows to interpret this component as the mean of the latent volatility process, which is only altered by jumps with permanent effect. The transitory effects are given by the mean reverting process $h_{i,t}$. Individual asset sensibility to the permanent volatility component is given by the scaling factor s_i^v . Factor models of stochastic volatility have known identification issues related to observationally equivalent combinations of parameter values and arbitrary factor rotation schemes (Geweke and Zhou (1996)). In order to circumvent these issues, we follow a traditional approach (Aguilar and West (2000), Chib, Nardari and Shephard (2006), Han (2005), Lopes and Carvalho (2007), Zhou, Nakajima and West (2014)) and set the first loading s_1^v (related to Bitcoin) to 1.

A second source of common discontinuous jumps affects return processes. These jumps capture events that jointly impact returns in the system. Differently from jumps

to global average volatility, which are permanent, jumps to return processes have contemporaneous effects only. If the realization of Bernoulli random variable δ_t^m , with parameter p_m , is 1, then $\varepsilon_{i,t}$ is sampled from a normal distribution with mean $\gamma_{i,t}$ and standard deviation 1 (a jump occurs); if δ_t^m comes up 0, then $\varepsilon_{i,t}$ is sampled from a standard normal distribution (no jump occurs). Note the model assumes distinct Bernoulli processes driving mean and variance jumps, thus jumps to returns and volatility are not correlated and can occur either simultaneously or independently. Individual sensibility of each asset to a jump in mean described by the scaling factor s_i^m . Here we also set s_1^m to one, as previous comments on identification apply. The intensity of jumps associated to s_1^m is given by the parameter σ_ν . The size of the jump at time t , z_t^m , follows a standard normal distribution.

Covariance between two individual assets i and j at time t is given by the coefficient $h_{(i,t),y}$, which depends on the scaling factors of sensibility to volatility jumps, and the permanent volatility component μ_t . In this model, covariance between assets is given by the exposition to the common variance component factor μ_t , since idiosyncratic transitory component $h_{i,t}$ are assumed to be independent between assets.

Due to the presence of several latent components, the estimation model (3.1)-(3.6) is based on a Bayesian mechanism using a mixed Markov Chain Monte Carlo procedure proposed by Laurini, Mauad and Auipe (2016). Jump processes δ_t^m and δ_t^v are sampled using the data augmenting scheme proposed by Albert and Chib (1993). A jump occurs if the realization of an auxiliary independent, uniformly distributed variable, exceeds some threshold — which is determined by unconditional jump probabilities p_m and p_v . Remaining latent processes are sampled though a mixture of Gibbs and Metropolis steps using the Slice Sampler of Neal (2003). This setup allows to avoid, for example, the linearization step of Qu and Perron (2013), or the mixture of Gaussians used by Kim et al (1998).

Estimation was carried out with 20,000 iterations of the sampling algorithm. In order to reduce the influence of initial values, the first 5,000 draws were discarded².

3.4 Results

Here we present and discuss parameter estimates of model (3.1)-(3.6) and comment on the implied path of latent variables. Then posterior means are employed as calibrated values for simulation exercises.

² Programs are available upon request.

3.4.1 Posterior estimates

Posterior descriptive statistics of parameters common to all assets in the double jump model (3.1)-(3.6) are presented in Table 3.5. Standard deviation of permanent volatility jumps σ^v is estimated between 0.617 and 1.337 with 95% probability, and the unconditional probability such jumps, p_v has posterior mean of 0.010, which implies an average of about 3.6 jumps per year. Return jumps are more frequent, with unconditional probability p_m estimated around 0.358, and standard deviation σ^m between 0.110 and 0.127 with 95% probability.

Table 3.5 – Posterior distributions common parameters

	Mean	Std.	2.5q	50q	97.5q	Skew.	Kurt
σ^v	0.910	0.186	0.617	0.885	1.337	0.636	3.245
σ^m	0.118	0.004	0.110	0.118	0.127	0.057	3.019
p_v	0.010	0.004	0.004	0.010	0.017	0.390	3.017
p_m	0.358	0.024	0.312	0.359	0.398	-0.286	2.339

Note: table summarizes posterior descriptive statistics of parameters of the multivariate jump model (3.1)-(3.6) common to all assets. The standard deviation of permanent volatility innovations σ^v has prior $IG(1.5, 1.25)$, the standard deviation of jumps to contemporaneous mean returns σ^m has prior $IG(0.5, 2.5)$, unconditional probabilities of jumps to volatility and mean, p_v and p_m , have prior $B(1, 40)$. Sampling was carried out with 20,000 repetitions, and statistics calculated after a burn-in period of 5,000 draws.

Table 3.6 summarizes posterior descriptive statistics of parameters in model (3.1)-(3.6) that are particular to each individual asset i : loadings associated to volatility and returns, s_i^v and s_i^m ; persistence parameter of autoregressive transitory volatility ϕ_i , and standard deviation of idiosyncratic volatility innovations σ_i^h . Ethereum and Bitcoin display higher persistence of transitory volatility h_t^i , suggesting these preeminent cryptocurrencies have (relatively) more autonomous volatility dynamics. Credibility intervals of persistence parameters ϕ_i are between 0.656 and 0.957, indicating some persistence of idiosyncratic volatility processes but no clear suggestion of nonstationarity. Volatility loadings s_i^v measure the exposure of each asset to the global time varying average volatility μ_t . Because μ_t is expressed in terms of log-volatility, smaller s_i^v means an increase in μ_t (in that it becomes less negative) will contribute less to the conditional volatility of asset i . That said, loadings s_i^v are all less than one, which is expected since Bitcoin is the less volatile cryptocurrency. Mean return jumps have amplified impact on altcoins other than Ripple; it is, their loadings s_i^m are above 1. It is the combination of jump probability p_m , standard deviation σ^m , and exposure s_i^m that determine the impact of mean return jumps on individual conditional volatility.

Figure 3.2 plots the common permanent volatility component μ_t in panel (a), and the probability of a permanent volatility jump in panel (b). From the beginning of our sample on August 16, 2015, until March 2017, log-variance component μ_t remains stable

Table 3.6 – Posterior distributions individual parameters

		Mean	Std.	2.5q	50q	97.5q	Skew.	Kurt
Bitcoin	s_1^v	1.000	0.000	1.000	1.000	1.000	0.000	0.000
	s_1^m	1.000	0.000	1.000	1.000	1.000	0.000	0.000
	ϕ_1	0.904	0.031	0.837	0.906	0.957	-0.433	3.139
	σ_1^h	0.306	0.091	0.147	0.305	0.488	0.221	2.682
Ethereum	s_2^v	0.840	0.037	0.761	0.841	0.905	-0.225	2.585
	s_2^m	1.132	0.032	1.071	1.132	1.192	-0.014	2.921
	ϕ_2	0.916	0.023	0.867	0.918	0.957	-0.421	3.224
	σ_2^h	0.324	0.081	0.188	0.319	0.496	0.386	2.728
Ripple	s_3^v	0.917	0.028	0.856	0.917	0.974	-0.097	3.159
	s_3^m	0.993	0.033	0.927	0.993	1.059	-0.001	2.957
	ϕ_3	0.770	0.038	0.694	0.771	0.842	-0.167	3.383
	σ_3^h	1.124	0.178	0.787	1.120	1.481	0.135	3.127
Litecoin	s_4^v	0.990	0.032	0.929	0.989	1.056	0.135	2.911
	s_4^m	1.090	0.031	1.028	1.090	1.152	0.024	3.117
	ϕ_4	0.766	0.052	0.656	0.768	0.860	-0.337	3.127
	σ_4^h	1.120	0.225	0.709	1.096	1.605	0.387	2.958
Stellar	s_5^v	0.818	0.028	0.770	0.815	0.879	0.440	2.889
	s_5^m	1.161	0.043	1.077	1.161	1.247	0.089	2.926
	ϕ_5	0.753	0.042	0.661	0.756	0.827	-0.442	3.094
	σ_5^h	0.771	0.134	0.544	0.756	1.076	0.538	3.483
Dash	s_6^v	0.865	0.026	0.823	0.862	0.920	0.492	2.630
	s_6^m	1.112	0.036	1.041	1.112	1.183	-0.037	2.971
	ϕ_6	0.785	0.046	0.681	0.790	0.864	-0.526	3.452
	σ_6^h	0.487	0.109	0.293	0.481	0.722	0.324	2.916
Monero	s_7^v	0.817	0.028	0.772	0.814	0.879	0.570	2.924
	s_7^m	1.192	0.039	1.116	1.192	1.269	0.015	2.913
	ϕ_7	0.867	0.029	0.804	0.869	0.916	-0.447	3.109
	σ_7^h	0.375	0.081	0.235	0.365	0.558	0.503	3.128
NEM	s_8^v	0.782	0.028	0.733	0.781	0.835	0.128	2.374
	s_8^m	1.267	0.040	1.192	1.265	1.348	0.138	2.988
	ϕ_8	0.814	0.042	0.721	0.819	0.886	-0.522	3.278
	σ_8^h	0.600	0.133	0.389	0.584	0.884	0.522	2.899
Verge	s_9^v	0.618	0.025	0.569	0.620	0.662	-0.175	3.149
	s_9^m	1.486	0.054	1.378	1.487	1.592	-0.047	3.000
	ϕ_9	0.828	0.043	0.736	0.831	0.905	-0.346	2.866
	σ_9^h	0.930	0.278	0.498	0.898	1.634	0.768	3.588

Note: table summarizes posterior descriptive statistics of parameters of the multivariate jump model (3.1)-(3.6) common to all assets. The standard deviation of permanent volatility innovations σ^v has prior $IG(1.5, 1.25)$, the standard deviation of jumps to contemporaneous mean returns σ^m has prior $IG(0.5, 2.5)$, unconditional probabilities of jumps to volatility and mean, p_v and p_m , have prior $B(1, 40)$. Sampling was carried out with 20,000 repetitions, and statistics calculated after a burn-in period of 5,000 draws.

Figure 3.2 – Volatility component and probability of jumps

Note: figure presents the permanent volatility component μ_t in panel (a) and the probability of a volatility jump in panel (b).

at about -8, which implies $\exp(-8/2) = 1.83\%$ daily return volatility in the long run. On March 9 2017, μ_t jumps to -6.7, implying 3.50%. This (first) high volatility regime is sustained until June 1st, when daily volatility falls to about 2.87%, still above pre-2017 levels. By the end of November 2017, amid soaring prices and great popular interest, daily average cryptocurrency volatility again rises to high levels — where it remains until mid-January 2018. From January 2018 through the end of our sample volatility steadily falls. This common long run average volatility path is consistent with what is found when one estimates the jump model for Bitcoin alone (see Chaim and Laurini (2018)).

In Appendix A we compare the common time varying average component μ_t just estimated with the univariate version of the model. The main difference from model (3.1)-(3.6) is that now the permanent component is no longer a common factor among all assets, and as the model is estimated individually for each asset all loading's parameters are fixed in the unit value. One can notice a spike in volatility during 2017 and sharp volatility

Figure 3.3 – Mean jumps and probability of jumps

Note: figure presents common return jumps γ_t in panel (a) and the probability of return jumps in panel (b).

reduction towards the end of the sample are features shared by most cryptocurrencies.

Jumps in mean returns are very frequent, as we can see in Figure 3.3. Unconditional probability p_m is estimated around 0.35 (see Table 3.5). This relatively high probability of joint return jumps suggests strong interconnection between cryptocurrencies, as already hinted by contemporaneous correlations in Table 3.3. Return jumps seem of larger magnitude and to happen more often from early 2017 onward. This phenomenon is evident from jumps $\gamma_{1,t}$ in panel (a) and probabilities of return jumps in panel (b), both from Figure 3.3³. The year of 2017 was a period of high returns and volatility for cryptocurrency markets. Indeed, we observe the global volatility component μ_t increases steeply in early 2017, then again by the end of the year, to then fall from early 2018 towards the end of our sample. This change in cryptocurrency volatility dynamics could be due to increased popular attention. Measures such as Google Trends Index point to a peak

³ Remember these are return jumps assuming loading $s_i^m = 1$. If the loading is not equal to one, the jump is scaled by the loading value.

in popular interest in cryptocurrencies during the year of 2017. See Kristoufek (2013a), Kristoufek (2015), Yelowitz and Wilson (2015), for examples of discussions on the relation between Bitcoin and Google Trends Index.

In order to illustrate some fit properties of the model, we take absolute returns as proxy of true volatility and compare accuracy measures against standard univariate stochastic volatility (SV) models. To estimate SV models, we employ the algorithm of Kastner and Frühwirth-Schnatter (2014) to draw a single chain of 30,000 sample draws. From Table 3.7 one can see the multivariate double jump model has smaller mean errors and mean squared errors, but larger root mean squared errors for six out of nine cryptocurrencies. Interestingly, the double jump model seems to, on average, underestimate volatility. One could argue that from the point of view of risk management it is undesirable to underestimate volatility. Still mean errors of the multivariate jump model are considerably smaller in magnitude. This evidence suggests model (3.1)-(3.6) has fit properties comparable to traditional models such as standard SV. It is difficult to compare the adjustment results obtained in our estimations with other results found in the literature, since due to the large temporal variations in the empirical standards of cryptocurrencies between the several samples studied in these works.

Table 3.7 – Model Fit Models

	Mean Error			Root Mean Squared Error			Mean Absolute Error		
	Jump	SV	$\Delta\%$	Jump	SV	$\Delta\%$	Jump	SV	$\Delta\%$
BTC	-0.00453	0.0077	58.8%	0.027	0.0236	114.4%	0.0165	0.0173	95.4%
ETH	-0.00361	0.0123	29.3%	0.0407	0.0369	110.3%	0.0282	0.0288	97.9%
XRP	-0.00532	0.0107	49.7%	0.0408	0.0414	98.6%	0.0231	0.0252	91.7%
LTC	-0.00902	0.0102	88.4%	0.0358	0.0327	109.5%	0.0207	0.022	94.1%
XLM	-0.0053	0.013	40.8%	0.0441	0.0453	97.4%	0.0275	0.0315	87.3%
DASH	-0.00547	0.0109	50.2%	0.0368	0.0353	104.2%	0.0242	0.0261	92.7%
XMR	-0.00484	0.013	37.2%	0.044	0.0415	106.0%	0.0293	0.0308	95.1%
XEM	-0.00516	0.015	34.4%	0.0524	0.0518	101.2%	0.0332	0.0359	92.5%
XVG	0.00611	0.0303	20.2%	0.0805	0.0903	89.1%	0.0574	0.0664	86.4%

Note: table reports comparison of forecast accuracy measures between the multivariate double jump model (Jump) and univariate standard log-normal stochastic volatility models (SV). Columns $\Delta\%$ stands for Jump model errors as proportion of SV error measures. 30000 sample draws using the algorithm of Kastner and Frühwirth-Schnatter (2014) are employed in estimating SV models.

3.4.2 Simulation Evidence for the Presence of Long Memory Processes

As discussed in Section 3.1, there is a long literature reporting the presence of long memory effects in digital currency financial data. Long range dependence in cryptocurrency volatility is a well-documented feature. Bariviera et al. (2017) calculate Hurst exponents for Bitcoin returns volatility and argue there is evidence of self-similarity. Lah-

miri, Bekiros and Salvi (2018) provide evidence of long-range dependence in seven Bitcoin markets using a fractionally integrated GARCH model under several specifications of error distribution.

The possible presence of long memory is an important effect since this may indicate the presence of market inefficiency for these assets (e.g., Kristoufek (2018)), having important consequences for the predictability of these assets and in portfolio allocation procedures and risk management. However, these possible long-memory effects may be spurious statistical effects generated by the presence of structural breaks, changes in parameters and regimes, and thus an effect generated by the aggregation of several short-memory processes, as discussed by Diebold and Inoue (2001) and Lai and Xing (2006). Charfeddine and Maouchi (2018) employ several tests to argue cryptocurrency volatility has indeed true long memory rather than level shifts.

Perron and Qu (2010) showed S&P500 volatility dependence features are better described by a stationary model with level shifts (for example the stochastic volatility model with permanent jumps to volatility of Qu and Perron (2013)) rather than a true long memory process. Charfeddine and Maouchi (2018) employ several spectral-based estimators to argue that regarding major cryptocurrencies, volatility indeed display long memory characteristics. Our exercise follows Perron and Qu (2010) argumentation in showing stationary models augmented with levels shifts are often the source of long memory features in financial asset data, and that this could be that case for cryptocurrencies.

To verify if the structure of jumps in the mean and volatility given by the model presented in Section 3.3 can generate these long memory effects, we perform Monte Carlo simulation. We employ posterior (mean) estimates to calibrate simulations exercises to verify whether the model considered is able to produce the increased temporal persistence characteristic of cryptocurrency volatility, and identify which features of those models are responsible for apparent long memory.

We used four different specifications in the Monte Carlo experiments to study the effect of different jump components on the long memory parameter. The first specification consists of simulating a Gaussian iid process, which serves as a reference for a short memory process with no jumps in the mean or variance. The next specification assumes jumps only in the common volatility component, and in this specification we assume that the returns have zero mean with only jumps in the common variance component. The third specification assumes that the common component of volatility is constant in time ($\mu_t = \mu_{t-1}$), and we only have the jump process in the mean given by equation (3.4). The process μ_t assumes the value at time 1 estimated by the empirical model. The fourth specification assumes the complete model with jumps in the mean and variance structure. As mentioned, we use the posterior mean of the parameters estimated in the

model applied to the real data as parameters in these different specifications. For these different specifications, we generate 2,000 simulated samples, for sample sizes of 200, 500 and 1000 observations, and estimated the long memory parameter d using the GPH estimator for two different measures of conditional variance, the squared returns and the log of the squared returns. The squared returns are a measure closer to the estimation of GARCH models, which assume that the squared returns follows an ARMA process for short memory or ARFIMA for long memory, and the log squared returns is a proxy linked to the stochastic volatility model, since usual forms of estimation for this model use a discretization based on the log squared returns as the dependent variable in a state space model. Table 3.8 reports the GPH estimates for the squared returns, and the Table 3.9 for the log squared returns. Specification details are presented in Appendix.

Table 3.8 – GPH Estimation for the Squared Returns

Model	S. Size		Bitcoin	Ethereum	Ripple	Litecoin	Stellar	Dash	Monero	NEM	Verge
iid	200	mean d	0.0028	-0.0002	0.0003	0.0028	0.0019	-0.0008	0.0028	-0.0011	0.0006
		prop. rej.	0.058	0.065	0.0595	0.071	0.0605	0.0575	0.0545	0.062	0.057
	500	mean d	0.0013	0.0009	0.0005	-0.0015	3.4e-05	0.0002	-0.0023	0.0012	-0.0010
		prop. rej.	0.0545	0.061	0.0555	0.0495	0.062	0.045	0.058	0.0575	0.0495
	1000	mean d	-0.0011	5.97e-05	-0.0017	-0.0020	-0.0005	0.00021	0.0014	-0.0022	0.00151
		prop. rej.	0.0525	0.051	0.051	0.049	0.0515	0.059	0.0475	0.0585	0.051
variance only	200	mean d	0.183	0.198	0.169	0.175	0.162	0.142	0.174	0.181	0.208
		prop. rej.	0.397	0.431	0.423	0.438	0.362	0.274	0.362	0.401	0.518
	500	mean d	0.195	0.219	0.136	0.141	0.146	0.133	0.182	0.168	0.185
		prop. rej.	0.658	0.755	0.48	0.475	0.467	0.401	0.617	0.56	0.645
	1000	mean d	- 0.202	0.224	0.107	0.118	0.124	0.124	0.183	0.151	0.166
		prop. rej.	0.879	0.927	0.452	0.517	0.534	0.521	0.819	0.658	0.749
mean only	200	mean d	0.144	0.188	0.162	0.172	0.155	0.127	0.172	0.178	0.209
		prop. rej.	0.289	0.402	0.408	0.419	0.338	0.237	0.356	0.386	0.52
	500	mean d	0.166	0.212	0.135	0.142	0.134	0.123	0.177	0.164	0.185
		prop. rej.	0.529	0.722	0.469	0.481	0.435	0.365	0.604	0.556	0.654
	1000	mean d	0.169	0.216	0.115	0.113	0.121	0.116	0.177	0.148	0.165
		prop. rej.	0.745	0.907	0.492	0.483	0.51	0.482	0.793	0.665	0.746
full model	200	mean d	0.145	0.19	0.162	0.165	0.154	0.125	0.171	0.171	0.212
		prop. rej.	0.293	0.419	0.41	0.397	0.34	0.241	0.348	0.376	0.528
	500	mean d	0.16	0.208	0.137	0.139	0.138	0.126	0.175	0.161	0.191
		prop. rej.	0.517	0.725	0.467	0.477	0.449	0.38	0.585	0.527	0.663
	1000	mean d	0.171	0.216	0.111	0.115	0.124	0.116	0.175	0.149	0.167
		prop. rej.	0.756	0.908	0.489	0.499	0.527	0.48	0.785	0.67	0.725

Table reports average d GPH estimates (using the $m = [T^{0.5}]$ lowest frequencies) and the proportion of rejection of the $d = 0$ hypothesis with 95% confidence level, from 2,000 squared returns simulations of four model specifications described in Appendix.

Table 3.9 – GPH Estimation for the Log of Squared Returns

Model	S. Size		Bitcoin	Ethereum	Ripple	Litecoin	Stellar	Dash	Monero	NEM	Verge
iid	200	mean d	0.0015	-0.0056	0.0034	-0.00012	-0.0016	-0.0009	-0.0016	0.0004	0.0012
		prop. rej.	0.0555	0.0645	0.069	0.0645	0.067	0.057	0.0545	0.055	0.0655
	500	mean d	-0.0017	-0.0004	-1.96e-05	0.0013	0.0014	-0.0003	0.0012	-0.0004	-0.0028
		prop. rej.	0.051	0.0555	0.058	0.054	0.058	0.0635	0.058	0.0535	0.0575
	1000	mean d	0.0006	0.0022	-0.0008	0.0003	7.55e-05	-5.73e-05	1.82e-05	-0.0011	-0.0011
		prop. rej.	0.062	0.058	0.054	0.0535	0.053	0.05	0.0565	0.0475	0.048
jumps in variance only	200	mean d	0.121	0.145	0.27	0.257	0.168	0.1	0.13	0.165	0.271
		prop. rej.	0.205	0.261	0.653	0.6	0.343	0.156	0.22	0.314	0.648
	500	mean d	0.133	0.169	0.239	0.231	0.156	0.0963	0.136	0.157	0.257
		prop. rej.	0.398	0.573	0.835	0.821	0.51	0.234	0.405	0.497	0.879
	1000	mean d	0.143	0.175	0.213	0.204	0.141	0.0926	0.136	0.147	0.247
		prop. rej.	0.659	0.807	0.92	0.898	0.645	0.333	0.605	0.674	0.973
jumps in mean only	200	mean d	0.0965	0.134	0.242	0.214	0.156	0.091	0.118	0.15	0.263
		prop. rej.	0.153	0.237	0.57	0.481	0.288	0.124	0.191	0.273	0.624
	500	mean d	0.105	0.156	0.215	0.195	0.143	0.0895	0.122	0.146	0.248
		prop. rej.	0.288	0.506	0.76	0.679	0.443	0.207	0.346	0.468	0.863
	1000	mean d	0.115	0.16	0.194	0.177	0.13	0.0824	0.123	0.135	0.239
		prop. rej.	0.478	0.746	0.87	0.812	0.57	0.283	0.525	0.6	0.966
jumps in mean and variance	200	mean d	0.0913	0.134	0.246	0.216	0.152	0.0893	0.113	0.151	0.263
		prop. rej.	0.148	0.227	0.58	0.485	0.289	0.141	0.179	0.271	0.642
	500	mean d	0.106	0.152	0.216	0.194	0.142	0.0855	0.12	0.144	0.254
		prop. rej.	0.289	0.497	0.76	0.69	0.447	0.188	0.336	0.453	0.88
	1000	mean d	0.114	0.161	0.198	0.179	0.13	0.0822	0.126	0.139	0.238
		prop. rej.	0.465	0.75	0.894	0.83	0.583	0.278	0.56	0.631	0.961

Table reports average d GPH estimates (using the $m = [T^{0.5}]$ lowest frequencies) and the proportion of rejection of the $d = 0$ hypothesis with 95% confidence level, from 2,000 log squared returns simulations of four model specifications described in Appendix.

In general, the results presented in Tables 3.8 and 3.9 are quite clear to indicate the effectiveness of jump components in generating long memory patterns for the measures of conditional variance studied. As expected, the iid process without jumps in the mean or variance presents estimated values for the coefficient d close to zero, with a proportion of rejections of the null hypothesis of short memory (d parameter equals zero) close to the expected level of 5%, confirming the good properties of the GPH estimator in the validity of null hypothesis. In all other specifications with the inclusion of jump components, the results are quite different. We can observe that both the individual inclusion of jump components in the mean or variance structures, as well as the inclusion of jump components in these two components at the same time, are able to generate long memory effects in the generated series.

We can observe that for specifications with jumps the estimated, coefficient d is close to the values found in the long memory estimates for cryptocurrencies (approximately ranging from 0.1 to 0.28 in our experiments), with a proportion of violations well above the expected 5% proportion of rejections, with this proportion increasing with the number of observations. These results are very important, since they show that different natures of jumps can induce the presence of long memory in the process of conditional variance. Jumps in the common component of volatility generate level changes in volatility, which are persistent and affect all series in the process, inducing a direct effect on the persistence of conditional volatility. This effect may be related to the influence of level changes and regime switching for the generation of long memory effects. But note that even a non-persistent process, such as the jump structure in the mean used in this work, can generate long memory effects for the conditional variance. This effect can be related to observations of long memory generated by the aggregation of distinct stochastic processes of short memory or zero memory. The combination of these two types of jumps also generates the long memory effects observed in financial assets, especially the cryptocurrency market studied in this work. Note that the results are robust to the choice of the squared returns or log squared returns and the sample size used.

3.4.3 Sub-sample Analysis

When looking at Figure 3.3 we can see that there is an apparent change in the process of jumps in the mean between the period until approximately the end of 2016 and the period after 2017. In this final period, we can note a significantly increase in the number of jumps in mean, which correspond to the common events directly affecting the returns on all assets.

To verify if there are relevant changes in the parameters of the model in these two periods, we estimate the model in two sub-samples. The first one runs from August 2015

until December 2016, and the second from January 2017 until October 2018. For space reasons we focus our analysis on the parameters related to the jumps in the mean, but the other results for this sub-sample analysis are available with the authors. The procedures of estimation and structure of priors are identical to those used in the estimation with the full sample.

The figure 3.4 shows the estimated posterior distribution for the two sub-samples for the parameter p_m , which is the unconditional probability of jumps in mean. We can note that there is a large difference between these two estimates. In the first sub-sample the parameter p_m is concentrated near zero, indicating that the common jumps in the returns have low probability, showing that until 2017 the chance of common events (jumps) in the returns is almost null. After 2017 this situation changes in a relevant way. We can observe that in the second sub-sample the posterior distribution of the parameter p_m concentrates around 0.52, indicating that there is now a relevant dynamic of direct transmission of shocks in returns for these assets.

The conditional jump probability in the mean for the two sub-periods confirms this interpretation, as shown in Figure 3.5. For the first sub-sample there is no day with jump probability greater than 0.3, while for the second sub-sample there is a relevant proportion of days with jump probability close to 1. This result indicates a relevant change in the cryptocurrency market from the year 2017. This result is consistent with the public interest in this market, which has grown substantially in the second sub-sample.

3.5 Conclusion

We investigated the joint dynamics of nine major cryptocurrencies by estimating a model with common discontinuous jumps to returns and volatility, introduced by Laurini, Mauad and Auibe (2016). This model is based on a nonlinear structure that captures the presence of common jumps (co-jumps) in the mean and conditional volatility structures, and allows to capture the most important stylized facts in the cryptocurrency market.

The structure of common jumps incorporates the existence of joint movements in the cryptocurrency market, a fact that is very important for the investors of these assets. Joint variations of high magnitude reflect changes in risk levels and transmission of shocks in this market. Unlike traditional assets such as shares traded on centralized stock exchanges, cryptocurrencies are subject to additional sources of risk such as regulatory changes, forks and possibilities for fraud and price manipulation via hacking, and to large effects on prices due to the possible existence of price bubbles in these assets. In general, the occurrence of these events has effects on the whole market of cryptocurrencies, generating the common movements observed in this market.

Figure 3.4 – Unconditional probability of jumps in mean - subsample analysis

[First subsample]

[Second subsample]

Note: Unconditional Probability of Jumps in Mean. (a) First subsample August 2015 - December 2016, (b) Second subsample - January 2017 - October 2018

The results obtained in the estimation of the model indicate that the common variance component μ_t remains relatively stable around -8 , implying 1.83% mean daily volatility from 2015-08 to 2017-01. One can notice two high volatility episodes, in early 2017 and then in early 2018, then steady reduction towards the end of our sample on October 2018. The path of common average volatility is similar to individually estimated ones. Transitory mean jumps occur, on average, once every three days; though become more frequent and of larger magnitude from 2017 onward.

These are some indications of a structural change in cryptocurrency dynamics in past years. From an estimation in sub-samples, it is possible to observe that from 2017 the return dynamics presents a high proportion of common shocks in the returns, indicating a substantial increase in the dependence of the returns in this market, with a high proportion of days where common shocks affect the returns on all assets.

Figure 3.5 – Probability of jumps in mean - subsample analysis
[First subsample]

[Second subsample]

Note: Probability of Jumps in Mean. First subsample August 2015 - December 2016, (b) Second subsample - January 2017 - October 2018

Another important result obtained in this work is to show that the observed long memory evidences for the most important cryptocurrencies can be generated by the presence of jumps in the mean and variance of these assets. The Monte Carlo experiments performed show that the individual jumping structures in the mean and variance can generate the observed long memory patterns, and thus giving an alternative interpretation to the reported market inefficiencies for these assets.

The methodology used in this work presents some important restrictions. The first restriction is that we consider all jump processes to be common among all assets, and all individual dynamics are placed in the specific factors, which are assumed to follow continuous random variables. One possibility is to introduce specific jump components to the individual assets, which would allow capturing specific large discrete variations. Although this possibility is interesting in terms of model fit, the number of latent parameters and

processes would be quite high, which can pose problems in terms of model estimation and possible overfitting. A second interesting modification is to allow the jump process to be time dependent, in contrast to the independent Bernoulli process used in this work. In a structure of dependent jumps the presence of a jump could increase or decrease the probability of a jump at the next moment. This possibility could be implemented using a temporal dependence structure for the parameters related to the probability of jumps. This is a possibility not yet explored in this class of models.

In this work our focus was to propose an identification mechanism for the effect of possible jumps on the mean and variance in the structure of returns and conditional volatilities for a relevant set of cryptocurrencies. The results indicate that these components are important in the temporal dynamics of these assets and can adequately explain the empirical characteristics of this sector, such as the presence of large price changes and persistent changes in volatility levels. We also show that this jumping mechanism can generate the observed long memory phenomena for this class of assets, and that this mechanism can be generated by common jumps in the mean, variance or in these two components simultaneously.

Important aspects in terms of portfolio allocation and risk management in the presence of these jump processes still need to be explored for this asset class. Using the same class of models used in this paper, Laurini, Mauad and Auibe (2016) show that jumps have important allocation and risk effects for companies in the oil sector. They show that a minimum variance allocation strategy based on the common component of conditional volatility has a better performance in terms of Sharpe ratio and turnover in relation to allocation strategies based on conditional volatility models of the GARCH family. They also show that the inclusion of jump components is important for the calculation of tail risk measures, such as Value at Risk. An important extension for this work would be to verify the performance of these methods for the cryptocurrency market, which presents complex challenges in terms of allocation and risk management.

Appendix

Simulation Details

Here we describe the Geweke and Porter-Hudak (1983) estimator and the specification of models employed in our simulation exercises.

GPH log-periodogram d estimator

Let $I_{x,T}(\omega_j)$ be the sample periodogram of a univariate time series $\{x_t\}_{t=0}^T$ at Fourier frequency $\omega_j = 2\pi j/T$ ($j = 1, \dots, [T/2]$), then the GPH estimate of d is given by

the least-squares regression

$$\log(I_{x,T}(\omega_j)) = c - 2d \log(2 \sin(\omega_j/w)) + \varepsilon_j, \quad (3.7)$$

using the $m = [T^a]$ frequencies closest to zero, $j = 1, \dots, m$. From a central limit theorem (see Beran (2017)) we have that

$$\sqrt{m}(\hat{d}_{GPH} - d) \rightarrow N\left(0, \frac{\pi^2}{24}\right). \quad (3.8)$$

Specification of simulated models

iid model: This is the simplest normal log-normal mixture model, with uncorrelated error terms.

$$y_t = \exp(h_t/2)\varepsilon_t, \quad (3.9)$$

$$\varepsilon_t \sim N(0, 1), \quad (3.10)$$

$$h_t \sim N(0, \sigma^2). \quad (3.11)$$

Where y_t are log-returns, h_t is an iid normal random variable with variance σ^2 , and ε_t is a standard normal disturbance.

Model jumps to variance only: this is the specification of Qu and Perron (2013), in which log-volatility is disentangled into a transitory, zero-mean, autoregressive term and a discontinuously evolving permanent average.

$$y_{i,t} = \exp\left(\frac{h_{i,t}}{2} + \frac{s_i^v \mu_t}{2}\right)\varepsilon_{i,t}, \quad \varepsilon_{i,t} \sim N(0, 1) \quad (3.12)$$

$$h_{i,t} = \phi_i h_{i,t-1} + \sigma_i^h z_{i,t}^h, \quad z_{i,t}^h \sim N(0, 1), \quad (3.13)$$

$$\mu_t = \mu_{t-1} + \delta_t^v \sigma^v z_t^v, \quad z_t^v \sim N(0, 1), \quad (3.14)$$

$$\delta_t^v \sim \text{Bernoulli}(p_v), \quad (3.15)$$

$$h_{(i,j),t} = s_i^v s_j^v \mu_t. \quad (3.16)$$

Where $y_{i,t}$ are log-returns of asset i , $h_{i,t}$ transitory autoregressive log-volatility, μ_t is the common, discontinuously varying, average log-volatility, s_i^v is the exposure of asset i to μ_t , p_v is the unconditional probability of a permanent jump to volatility, and $h_{(i,j),t}$ is the implied covariance between assets i and j at time t .

Model with jumps to mean only: This model is a modification of the multivariate double jump model employed by Laurini, Mauad and Auibe (2016), in which log-volatility follows a stationary process, and there are only common contemporaneous

jumps to mean returns:

$$y_{i,t} = \exp\left(\frac{h_{i,t}}{2} + \frac{s_i^v \mu_t}{2}\right) \varepsilon_{i,t}, \quad \varepsilon_{i,t} \sim N(\gamma_{i,t}, 1) \quad (3.17)$$

$$h_{i,t} = \phi_i h_{i,t-1} + \sigma_i^h z_{i,t}^h, \quad z_{i,t}^h \sim N(0, 1), \quad (3.18)$$

$$\mu_t = \mu_{t-1}, \quad (3.19)$$

$$\gamma_{i,t} = s_i^m \delta_t^m \sigma^m z_t^m, \quad z_t^m \sim N(0, 1), \quad (3.20)$$

$$\delta_t^m \sim \text{Bernoulli}(p_m), \quad (3.21)$$

$$h_{(i,j),t} = s_i^v s_j^v \mu_t. \quad (3.22)$$

where $y_{i,t}$ are log-returns of asset i , $h_{i,t}$ transitory autoregressive log-volatility, μ_t is the common, constant, average log-volatility, $\gamma_{i,t}$ is the return jump of asset i at time t , which depends on the loading s_i^m , the standard deviation σ^m , the standard normal error term z_t^m , and the Bernoulli random variable $\gamma_{i,t}$, which has unconditional probability of success p_m .

Univariate Jump Models

As robustness exercise, we estimate nine univariate versions of model (3.1)-(3.6), and compare the implied permanent volatility components between multivariate and univariate specifications. In the univariate version of the jumps model, the μ_t component loses the common factor in volatility interpretation, but still represents permanent changes in the level of volatility. In this formulation all load parameters are equal to one, and there is no dynamics for the covariance included in the model. This model is a generalization of the models used in Qu and Perron (2013) and Laurini and Mauad (2014), including the jump component in mean. We focused on this section in comparing the long-term component μ_t obtained in the multivariate estimation with the values obtained in the univariate estimates. For space reasons we do not show the estimated parameters in the univariate versions and the other results, but they are available with the authors.

In Figure 3.6 we plot the permanent volatility component μ_t of the multivariate double jump model (3.1)-(3.6) against the time-varying volatility obtained when the same model is estimated for each cryptocurrency individually.

The results obtained show that in general the permanent component estimated in the multivariate and univariate versions of the jump model has similar patterns, especially in terms of average values. As discussed in Section 3.4, one can notice a spike in volatility during 2017 and sharp volatility reduction from early 2018 towards the end of the sample. This feature is shared by most cryptocurrencies, indicating the global volatility component is a good representation of the volatility level in cryptocurrency markets. The μ_t component estimated by the multivariate model appears to be an aggregation of

Figure 3.6 – Common and individual long-run volatility

Note: figure 3.6 plots the common time-varying average volatility μ_t estimated in Section 4 (solid blue lines) against the implied trend volatility when model (3.1)-(3.6) is estimated for each cryptocurrency individually (solid black lines).

the permanent components estimated by the univariate models, showing a consistency between the results of these estimates, showing the robustness of the results obtained by this model.

4 Is Bitcoin a bubble?

Pedro Chaim
Márcio P. Laurini

Abstract

The narrative of a Bitcoin is a bubble is very common. We employ statistical techniques to empirically evaluate such claim. A branch of literature links the existence of a bubble in some financial asset's price to strict local martingales — a finitely lived asset has a bubble if, and only if, it is a strict local martingale under the equivalent risk-neutral measure. A diffusion process is a strict local martingale if its volatility increases faster than linearly as its level grows. We apply a nonparametric method to estimate the volatility function of Bitcoin daily and high frequency prices, as well as of more traditional financial assets. We then estimate the stochastic volatility model of Andersen and Piterbarg (2007), whose parameter space has a specific subset under which the asset's price is a strict local martingale. Results suggest the existence of a bubble in Bitcoin prices from early 2013 to mid 2014, but, interestingly, not in late 2017.

Keywords: Bitcoin; Cryptocurrencies; Financial Bubbles; Strict Local Martingales.

4.1 Introduction

The surge in Bitcoin and other cryptocurrencies prices in recent years has attracted remarkable media and academic attention. A common narrative is that Bitcoin is undergoing a price bubble (Krugman (2018), Bloomberg (2017)). We employ statistical techniques to empirically evaluate such claim.

Intuitively, an asset undergoing bubble pricing has market price that exceeds the price a rational agent would pay for it — its “fundamental value”. The fundamental price of a financial asset is often thought as the price a risk-neutral rational agent would pay for it, and calculated by discounting expected future cash flows from the asset. The existence of a bubble in Bitcoin's prices can be justified by several factors. The first is the presence of exaggerated expectations about the adoption of Bitcoin (and cryptocurrencies in general) as a practical mean of payment substitution of traditional money assets. A second possible motivation is the cheapening of mining technologies in conjunction with higher mining difficulty. For example, Kristoufek (2015) argues mining difficulty measures are positively correlated with price, and that there is a feedback relation in which higher prices and

higher popular attention lead to more miners. Another possible interpretation is related to convenience yield explanations of the Dot-com (internet) bubble in late 1990s and early 2000s (Cochrane (2002)), and also feedback mechanisms between Internet searches on virtual currencies and the prices of these assets (e.g., Kristoufek (2013b) and Yelowitz and Wilson (2015)).

Camerer (1989) reviews early attempts of identifying financial bubbles which involves proposing a model for some asset’s fundamental price. Defining an asset’s fundamental price is problematic since future cash flows are not observed (Obayashi, Protter and Yang (2017)). Because in those tests one has to specify a model for the fundamental value, discrepancy between the market price and the fundamental price can be either due to the presence of a bubble, or due to model misspecification. A critical issue with this joint hypothesis is that the model for the fundamental value cannot be independently validated. The theory outlined by Jarrow, Protter and Shimbo (2007), Jarrow, Protter and Shimbo (2010) and Jarrow, Kchia and Protter (2011a) might prove helpful. It links the presence of a bubble in some asset price S to the nonexistence of an equivalent risk-neutral martingale measure. The key point is that there exists a bubble if and only if S is not a “true” martingale, but a strict local martingale under the equivalent risk-neutral probability measure.

We explore two empirical approaches. In the first one we estimate the volatility function $\sigma^2(\cdot)$ of the diffusion process driving S . A testable condition (Condition (4.7) detailed in Section 2) states that S has a price bubble (i.e, is a strict local martingale) if its volatility $\sigma^2(\cdot)$ grows faster than linearly with higher price level S_t . The second one involves specifying a stochastic volatility model for the risky asset price S_t for which the parameter space can be partitioned in two mutually exclusive sets, one where the process exhibits a bubble and one where it does not (this draws on results of Andersen and Piterbarg (2007)). Because market prices S_t are observed, the estimated model can be independently validated.

We analyze Bitcoin daily¹ and high-frequency five minute prices between January 2013 and September 2018. Motivated by major developments in Bitcoin markets, we split the data in three subsamples. The first goes from January 1st 2013 to March 31 2014. This period is marked by the Mt. Gox affair, it covers the first time Bitcoin crossed the \$1000 mark (by the end of 2013), then fell around \$400. The second subsample ranges from April 1st 2014 to December 31 2016; this is a period of (relatively) low volatility and average returns in cryptocurrency markets. Volatility functions are estimated for daily and high-frequency data². In order to provide comparison context we repeat the exercises

¹ An average across major Bitcoin exchanges calculated by blockchain.com. Prices are collected at 22:00 or 21:00. We chose to leave out early periods due to low market liquidity.

² We do not estimate the stochastic volatility model with high-frequency data because the number of

done on Bitcoin for more “traditional” financial assets: the S&P500 index, the euro-dollar exchange rate, gold dollar prices, and Brent oil prices.

Results from both nonparametric estimation of volatility functions and Bayesian estimation of Andersen and Piterbarg (2007) model point to the presence of a Bitcoin price bubble in the first subsample — price volatility function is increasing for higher observed prices, and parameter estimates from Andersen and Piterbarg (2007) model fall in the strict local martingale range. Interestingly, the 2017-2018 period, although very volatile, is not characterized as a bubble following our criteria.

Statistical analysis of cryptocurrency markets is a rapidly expanding research field. Literature seems specially keen on market efficiency questions, and potential price (and volatility) “drivers”. Urquhart (2016) employ temporal dependence tests to Bitcoin returns to evaluate the weak market efficiency hypothesis. He finds that as the Bitcoin market grows bigger with time, it also appears more informatinally efficient. Ciaian, Rajcaniova and Kancs (2016) build on the gold standard price level framework of Barro (1979) to formulate and test hypotheses regarding the determination of Bitcoin prices; they find supply and demand forces and investor attractiveness are relevant in determining Bitcoin prices, but macrofinancial developments do not impact it. Garcia et al. (2014) employ multivariate autorregressive models to study the relationship between Bitcoin prices and what they call “digital socio-economic signals”. They find a feedback relation between Bitcoin (differenced) prices and volume of word-of-mouth communication, and volume of new adopters³.

There is some literature regarding bubbles in Bitcoin prices. Cheah and Fry (2015) analyze daily bitcoin prices from 2011 to 2014. They argue bitcoin’s sharp rise in late 2013, followed by a bust was a bubble. This is the same period in which price manipulation occurred, as documented by Gandal et al. (2018). MacDonell (2014) fits a log-periodic law model to bitcoin prices, and argues it ex-ante predicts the late 2013 crash. Corbet, Lucey and Yarovaya (2018) relate Bitcoin and Ethereum prices to “fundamental drivers” to date several bubble periods, including late 2013 and the second semester of 2017. Cheung, Roca and Su (2015), Su et al. (2018), Bouri, Shahzad and Roubaud (2018) apply the methodology of Phillips, Shi and Yu (2015), which is based on rolling window Augmented Dickey-Fuller tests, to datestamp bubbles in Bitcoin markets. Several periods of “explosivity” are highlighted, specially through 2017.

The remainder of this paper is structured as follows. Section 4.2 reviews the relation between financial bubbles and strict local martingale, and discusses key results and empirical exercises in this literature. Section 4.3 presents and discusses our results. Section 4.4 concludes.

observations makes Bayesian posterior simulation prohibitively expensive.

³ Authors argue this leads to “price bubble in absence of exogenous stimuli”.

4.2 Asset Price Bubbles and Strict Local Martingales

It is very difficult to say exactly if a bubble is present in the value of an asset. An informal definition of a bubble is the existence of a systematic deviation of the market price from the fair price of that asset, corresponding to the net present value of the future cash flows from that asset. A first difficulty in evaluating a possible bubble is that future cash flows are uncertain, so this assessment requires forecasting these values for the entire future time horizon⁴.

This forecast stage usually involves a great deal of subjectivity, since it involves all possible scenarios in the future. A second difficulty in this calculation is the discount rate to be used, also requiring a forecast for the entire future horizon of the asset. Normally these two steps are evaluated jointly, through risk premium on the discount rate as protection of the uncertainty present in the valuation, thus incorporating a risk-return relationship in the pricing of this asset. Note that this does not completely resolve the great uncertainty in the assessment of the fair value of this asset. Usually the existence of a bubble in the price of an asset is linked to the presence of exacerbated expectations about future cash flows, and bubble rupture may occur when there is a disappointment about the expected values of these future yields. Note that there is an additional difficulty in evaluating the presence of a bubble in the value of Bitcoin, since in this case there are no direct cash flows associated with this asset, such as payment of dividends in the case of equity shares. Thus the present value of this asset depends on the expectations about future capital gains (price appreciation), which makes this valuation even more difficult. See for example Jarrow, Kchia and Protter (2011a) for a discussion about this feature.

It is possible to notice that this form of evaluation is based on unobserved elements (Obayashi, Protter and Yang (2017)), and thus the definition of the presence of a bubble is not consensual, since different discount rates or expectations of future cash flows or price gains can lead to different probabilities for the probability of a bubble. Camerer (1989) argues this identification problem makes so that this traditional empirical bubble detection literature is inconclusive, but also suggests the issue can be circumvented if bubbles have common statistical properties that do not depend on the specification of market fundamentals that drives intrinsic value.

To avoid this possible ambiguity, an alternative is to use the so-called no-arbitrage valuation, which provides a direct way to check if an asset is correctly priced. In no-arbitrage theory an asset is correctly priced if the pricing process is consistent with a stochastic process known as martingale after adjusting for financing costs and risk, which is mathematically accomplished through a change of measure (the equivalent martingale

⁴ See Flood and Hodrick (1986), Smith, Suchanek and Williams (1988), Camerer and Weigelt (1991) for examples of this literature.

measure). A martingale is a process whose forecast for all future horizons of time is given by the current value of the asset, and thus a simple way to test for the correct pricing is to check the consistency of the price observed with a martingale process. The presence of a bubble imposes some additional mathematical difficulties in the assessment by no-arbitrage, which requires some additional characterizations. The no-arbitrage theory characterizes the existence of a bubble through the presence of a process known as strict local martingale, and thus provides a testable condition for the presence of a bubble that circumvents the difficulties existing in the usual cash flow assessment procedures.

Let us formalize the idea of a bubble. Because the literature on the subject is exceedingly vast and diverse, we focus on the strict local martingale theory of financial bubbles developed by Jarrow, Protter and Shimbo (2007), Jarrow, Protter and Shimbo (2010), Jarrow, Kchia and Protter (2011a).

Following Protter (2013), consider finitely lived economy in which are traded a risky asset and a money market account in the time interval $[0, T]$. For simplicity, assume there are no cash flows associated with the asset. Consider a complete probability space $(\Omega, \mathcal{F}_t, P)$, and a filtration $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$. Let $r = (r_t)_{t \geq 0}$ denote the instantaneous default-free spot interest rate, then the money market account value at time t is

$$B_t = \exp \left(\int_0^t r_u du \right). \quad (4.1)$$

No arbitrage in the sense of No Free Lunch with Vanishing Risk (NFLVR) implies that there exists a probability measure Q , equivalent to the physical probability measure P , such that

$$\frac{S_t}{B_t} \quad (4.2)$$

is a local martingale under Q . This is the first theorem of asset pricing, as stated by Delbaen and Schachermayer (1998).

A local martingale is an adapted càdlàg stochastic process M_t , for which there exists an increasing sequence of stopping times $(\tau_n)_{n \geq 0}$, such that $\lim_{n \rightarrow \infty} \tau_n = \infty$, and $M_{\min\{t, \tau_n\}}$ is a martingale with probability one for all n . Jarrow (2012) explains that a local martingale is a generalization of a martingale that extends the martingale's “fair game” property to a “game” that has a random termination time, which depends on information generated when “playing the game”.

Let S_T be the liquidation value of the asset at time T . Protter (2013) defines the fundamental price of the asset, S_t^* , which pays no dividends, as the expected value of the liquidation value S_T under the equivalent local martingale measure Q ,

$$S_t^* = E_Q \left(\frac{S_T}{B_T} \mathbf{1}_{\{t \leq T\}} \middle| \mathcal{F}_t \right) B_t. \quad (4.3)$$

The risky asset's fundamental value is equal to the price one would pay for it, if after the purchase, one had to hold the asset until time T . Protter (2013) shows S^* is a true martingale under Q .

The asset price bubble β_t is defined as the difference between the observed market price S_t and the fundamental value S_t^* . Such difference is assumed nonnegative.

$$\beta_t = S_t - S_t^* \geq 0. \quad (4.4)$$

Protter (2013) shows that saying some finitely-lived risky asset S_t has non-zero bubble β_t on $[0, T]$ is equivalent to say S_t is a strict Q -local martingale.

Jarrow, Kchia and Protter (2011a) discuss the intuition behind the distinction between a nonnegative martingale and a strict local martingale, $S \geq 0$, follows from the fact that S is always a supermartingale and is a true martingale if and only if it has constant expectation. Then if S is a strict local martingale its expectation decreases over time. The typical behavior of a strict local martingale process is to shoot up to high values, to then drop down and stay low.

Assume the risky asset price $S = (S_t)_{t \geq 0}$ is driven by a standard stochastic differential equation

$$dS_t = \mu(S_t)dt + \sigma(S_t)dW_t^S. \quad (4.5)$$

NFLVR implies that there exists an equivalent local martingale measure under which the SDE (4.5) simplifies to

$$S_t = S_0 + \int_0^t \sigma(S_u)dW_u^S. \quad (4.6)$$

It is shown by Kotani (2006) and Mijatović and Urusov (2012) that S is a strict local martingale if, and only if

$$\int_\delta^\infty \frac{S_t}{\sigma^2(S_t)} dS_t < \infty, \quad (4.7)$$

for all $\delta > 0$. Jarrow (2012) notes condition (4.7) states that an asset is a strict local martingale if and only if its volatility, as function of its price, increases faster than linearly. As Jarrow, Kchia and Protter (2011b) explain, this theory tells us that if the graph of the volatility versus the asset price tends to infinity at a faster rate than does the graph of $f(x) = x$, then there is a bubble.

The integral (4.7) can be empirically evaluated by estimating the diffusion coefficient, or volatility function, $\sigma^2(\cdot)$. Whether S is a true Q -martingale or a strict Q -local martingale will depend on the behavior of the estimated volatility $\hat{\sigma}^2(\cdot)$ as $S_t \rightarrow \infty$.

The volatility function $\sigma^2(\cdot)$ can be estimated by kernel-based nonparametric estimators, such as the one introduced by Florens-Zmirou (1993) and Jiang and Knight (1997). This approach is appealing because little structure is imposed on the dynamics of S , but is hindered by the fact that one can only compute $\hat{\sigma}(S_t)$ for observed values of S .

Jarrow, Kchia and Protter (2011a) introduce an alternative method of estimation for the volatility function based on a parameterized family of Reproducing Kernel Hilbert Spaces, which allows for the extrapolation of $\hat{\sigma}^2(x)$ for values x not visited by the sample path of process S . Jarrow, Kchia and Protter (2011b) employ such method to analyze LinkedIn stock prices through May 2011, they infer the presence of a bubble. The same method is applied by Jarrow (2016) to gold prices, no bubble is detected.

Now consider the dynamics of S follow the stochastic volatility model studied by Andersen and Piterbarg (2007).

$$\begin{aligned} dS_t &= \lambda S_t \sqrt{V_t} dW_t^S, \\ dV_t &= \kappa(\theta - V_t) + \epsilon V_t^p dW_t^V, \end{aligned} \tag{4.8}$$

where $\lambda, \kappa, \theta, p > 0$, and (W_t^S, W_t^V) is a two dimensional Brownian motion with correlation coefficient ρ .

Results of Andersen and Piterbarg (2007) and Protter (2013) describe how the parameter space of model (4.8) can be divided in two disjoint sets, such that a specific parameter range characterizes S as a strict local martingale; thus allowing for a testable implication of the martingale hypothesis.

Theorem 1. *For model (4.8): if $\rho \leq 0$, S is a true martingale; if $\rho > 0$ and $p \leq 1/2$ or $p > 3/2$, S is a true martingale; if $\rho > 0$ and $1/2 < p < 3/2$, then S is a strict local martingale. For the case $p = 3/2$, S is a true martingale if $\rho \leq 1/2\epsilon\lambda^{-1}$, and S is a strict local martingale if $\rho > \epsilon\lambda^{-1}/2$.*

Proof. *See Proposition 2.5 of Andersen and Piterbarg (2007).*

Baldeaux, Ignatieva and Platen (2018) investigate the presence of bubbles in money markets by estimating models similar to (4.8), and comparing estimated parameter values to the conditions in Theorem 1. They find the fitted models can produce bubbles.

4.3 Empirical Evidence

4.3.1 Descriptive Statistics

Selected descriptive statistics of Bitcoin log-returns are displayed in Table 4.1. Motivated by major developments in Bitcoin markets and existing empirical evidence, we

split the data in three subsamples. The first goes from January 1st 2013 to March 31 2014. This period is marked by the Mt. Gox affair, it covers the first time Bitcoin crossed the \$1000 mark (by the end of 2013), then fell around \$400; this episode is identified by Cheah and Fry (2015) as a bubble. The second subsample ranges from April 1st 2014 to December 31 2016; this is a period of low volatility and average returns (relatively, see second line of Table 4.1) in cryptocurrency markets. The third starts on January 1st 2017 and goes until the end of our sample, on September 1st 2018 — it captures the remarkable attention cryptocurrencies received during 2017 and early 2018, and is dated as a bubble by Corbet, Lucey and Yarovaya (2018). End dates for subsamples were motivated by the conditional volatility level of Bitcoin returns when contrasted to unconditional average volatility (see Chaim and Laurini (2018)).

In order to provide some context, we repeat the exercises performed on Bitcoin for a few, more traditional, financial assets. We analyze the S&P 500 index, the euro-dollar exchange rate, Brent oil prices, and dollar gold prices. Descriptive statistics for log-returns of those assets are also shown in Table 1.

Table 4.1 – Descriptive statistics of daily log-returns

	Mean	Std.Error	Min.	Median	Max.	Skew.	Kurt.
BTC 2013-01 to 2014-03	0.775	7.105	-47.831	0.521	35.879	-1.049	11.083
BTC 2014-04 to 2016-12	0.053	3.214	-26.862	0.096	21.168	-0.377	10.698
BTC 2017-01 to 2018-09	0.326	4.771	-22.571	0.522	24.661	-0.255	3.267
BTC full sample	0.302	4.769	-47.831	0.246	35.879	-0.764	15.162
S&P 500	0.048	0.771	-4.184	0.057	3.829	-0.591	3.268
EUR-USD	-0.009	0.533	-2.672	-0.008	3.064	0.042	2.099
Gold	-0.023	0.903	-8.913	-0.025	4.788	-0.685	8.239
Oil	-0.027	1.972	-8.083	0.000	9.896	0.370	2.994

Note: table reports descriptive statistics of daily log-returns of Bitcoin, including three subsamples, the S&P500 index, the EUR-USD exchange rate, dollar gold prices, and Brent oil prices, for the sample 2013-01 - 2018-09. First column displays mean. Second column displays standard errors. Third and fifth columns show the smallest and largest observations, respectively. Fourth column displays medians.

Sixth column show skewness coefficients. Seventh column displays raw kurtosis coefficients.

Camerer (1989) argues intuitively that since bubbles will cause some extremely large positive price changes as they grow (especially during the last stages of bubble growth), and even larger negative price changes when they burst, the distribution of returns should have negative skewness and large kurtosis if a bubble exists. Notice from Table 1 that Bitcoin returns are leptokurtic, with kurtosis well above what one would expect of a Normal distribution, and have negative skewness. One can take this as some initial positive evidence towards the presence of price bubbles.

Table 2 shows descriptive statistics for high-frequency log-returns data of the Bit-

Figure 4.1 – Daily Bitcoin dollar prices and returns

Note: figure plots Bitcoin daily prices and returns from January 2nd 2013 to September 1st 2018.
Dashed vertical lines represent subsample endpoints.

coin series. These returns are constructed using a 5-minute price aggregation, in this case using the last five-minute value of the tick-by-tick operations in Bitcoin Bitstamp exchange. The use of high-frequency data is especially interesting for this asset since it is traded continuously, unlike traditional assets with fixed periods of operation. The use of high-frequency data allows us to analyze whether the bubble detection processes are sensitive to the analyzed frequency of the data. We used a 5-minute aggregation to have a regular frequency of data, and also to mitigate possible contaminations by microstructure noise that affect high-frequency estimates. We can see that important characteristics remain present at this frequency of observation. In particular it is possible to note the existence of extreme values of returns, asymmetry patterns and especially high kurtosis, which may be associated with the presence of bubbles, as discussed above.

Table 4.2 – Descriptive statistics of five minutes Bitcoin log-returns

	Mean	Std.Error	Min.	Median	Max.	Skew.	Kurt.
BTC 2013-01 to 2014-03	0.003	0.713	-22.019	0.000	55.675	4.853	451.039
BTC 2014-04 to 2016-12	0.000	0.261	-15.069	0.000	9.820	-0.944	95.758
BTC 2017-01 to 2018-09	0.001	0.357	-20.264	0.001	7.450	-1.264	88.963
BTC full sample	0.001	0.429	-22.019	0.000	55.675	4.769	798.267

Note: Table reports descriptive statistics of high frequency five minute returns of Bitcoin, including three subsamples. First column displays mean. Second column displays standard errors. Third and fifth columns show the smallest and largest observations, respectively. Fourth column displays medians.

Sixth column show skewness coefficients. Seventh column displays raw kurtosis coefficients.

4.3.2 Volatility Function

Recall from our discussion in Section 4.2 that condition (4.7) implies a continuous stochastic process $S = (S_t)_{t \geq 0}$ is a strict local martingale if, and only if, its volatility increases fast enough as its level grows. That is, if the integral below is finite.

$$\int_{\delta}^{\infty} \frac{S_t}{\sigma^2(S_t)} dS_t < \infty, \quad \forall \delta > 0.$$

We present estimates the volatility function $\sigma^2(\cdot)$ based on the nonparametric estimator of Florens-Zmirou (1993) and Jiang and Knight (1997).

Consider the continuous stochastic process $S = (S_t)_{0 \leq t \leq T}$, let $\{S_t = S_{t_1}, S_{t_2}, \dots, S_{t_n}\}$ be n equispaced observations of process S at times $\{t_1 = \Delta_n, t_2 = 2\Delta_n, \dots, t_n = n\Delta_n\}$, with $\Delta_n = T/n$. When the sampling path of the diffusion process S_t visits x ,

$$\hat{\sigma}^2(x) = \frac{\sum_{i=1}^{n-1} nK\left(\frac{S_{i\Delta_n} - x}{h_n}\right) [S_{(1+i)\Delta_n} - S_{i\Delta_n}]^2}{\sum_{i=1}^n TK\left(\frac{S_{i\Delta_n} - x}{h_n}\right)}, \quad (4.9)$$

where $h_n > 0$ is a bandwidth parameter, and $K(\cdot)$ is a kernel function. We employ a Gaussian kernel, and set the bandwidth parameter specified by Kullback-Leibler cross-validation (Hurvich, Simonoff and Tsai (1998)).

Inference on whether S is a true martingale or a strict local martingale comes from the properties of $\sigma^2(S_t)$ as $S_t \rightarrow \infty$. That is, if volatility grows faster than linearly as level increases, then the underlying process is a strict local martingale according to condition (4.7). A limitation of this nonparametric estimation approach is that we can only estimate $\sigma^2(x)$ for values of x visited by the sample path of S_t .

Figure 4.2 displays nonparametric estimates of Bitcoin daily volatility functions $\hat{\sigma}^2(S_t)$ (left panels), and the ratio $S_t/\hat{\sigma}^2(S_t)$, integrand of condition (4.7), for the three considered subsamples and the whole sample. Figure 4.3 shows these estimated quantities for high frequency Bitcoin prices, for each subsample. Because there are more observations

for each price level, high frequency estimates are more precise. We repeat this exercise for the S&P500 index, the USD-EUR exchange rate, dollar gold prices, and Brent oil prices in Figure 4.4. We experimented with log-prices instead of prices but results were similar and uninformative, thus we do not report here.

Top panels (a) present estimates for the first Bitcoin subsample (2013-01-01 to 2014-03-31). Daily volatility function in Figure 4.2 seems to increase with level, but confidence bands are broad, particularly for prices above \$900. Turning to high frequency estimates in panel (a) of Figure 4.3 we see a steep slope of the variance function for higher price levels. This suggests condition (4.8) holds and Bitcoin follows a strict local martingale during this period.

Mid panels (b) in Figure 4.2 display daily estimates for the second subsample (2014-04-01 to 2016-12-31). If we compare the magnitudes of vertical axes, it is clear that Bitcoin is much less volatile during this second subsample, confirming descriptive statistics in Table 4.1. Volatility function estimates from both daily and high frequency data stabilize for higher Bitcoin prices. Integrand $S_t/\hat{\sigma}^2(S_t)$ in right panel (b) is more similar to those from Oil and Gold (Figure 4.4) than to other Bitcoin subsamples. The results obtained for the traditional assets indicate the absence of bubble evidence in the period under analysis. This indicates Bitcoin do not display a bubble during this period.

Estimates for the third subsample (2017-01-01 to 2018-09-01) are displayed in mid panels (c). In Figure 4.2 we see daily estimates are monotonically increasing but confidence bands are very broad. Interestingly, high-frequency estimates in Figure 4.3 show volatility increases steeply until around the \$13,000 mark, remains stable until \$15,000, and clearly falls for higher price levels. We take this evidence as indication the late 2017 period was not a bubble according to condition (4.7).

4.3.3 Stochastic Volatility Model

We now estimate the stochastic volatility model of Andersen and Piterbarg (2007), discussed in Section 4.2.

$$\begin{aligned} dS_t &= \lambda S_t \sqrt{V_t} dW_t^S, \\ dV_t &= \kappa(\theta - V_t) + \epsilon V_t^p dW_t^V, \end{aligned}$$

where $\lambda, \kappa, \theta, p > 0$, and (W_t^S, W_t^V) is a two dimensional Brownian motion with correlation coefficient ρ . As described in Theorem 1, the parameter space of this model is divided in two disjoint sets: one under which S is a true martingale, and another under which S is a strict local martingale.

Due to the presence of the latent variable V_t we estimate model (4.8) by Bayesian posterior simulation methods, specifically, the Hamiltonian Monte Carlo (HMC) algorithm

Figure 4.2 – Variance function estimates - daily Bitcoin prices

Note: figure displays nonparametric estimations of variance functions $\sigma^2(S_t)$ (left panels) and relation $S_t/\sigma^2(S_t)$ (right panels) for the three considered subsamples and the full sample for daily Bitcoin prices. Prices are collected at 21:00 or 22:00. We employed the Florens-Zmirou (1993) estimator with a Gaussian kernel. Shaded area represents a 95% confidence band, obtained from 10000 bootstrap simulations.

Figure 4.3 – Variance function estimates - high-frequency Bitcoin prices

Note: figure displays nonparametric estimations of variance functions $\sigma^2(S_t)$ (left panels) and relation $S_t/\sigma^2(S_t)$ (right panels) for the three considered subsamples and the full sample for high-frequency Bitcoin prices, collected every 10 minutes. We employed the Florens-Zmirou (1993) estimator with a Gaussian kernel. Shaded area represents a 95% confidence band, obtained from 10000 bootstrap simulations.

Figure 4.4 – Variance function estimates - S&P500, EUR-USD, Gold, and Oil

Note: figure displays nonparametric estimations of variance functions $\sigma^2(S_t)$ (left panels) and relation $S_t/\sigma^2(S_t)$ (right panels) for four selected daily financial asset: the S&P500 index, the EUR-USD exchange rate, dollar gold prices, and Brent oil prices. We employed the Florens-Zmirou (1993) estimator with a Gaussian kernel. Shaded area represents a 95% confidence band, obtained from 10000 bootstrap simulations.

introduced by Duane et al. (1987)⁵. HMC augments the usual random walk posterior sampling scheme of the Metropolis-Hastings (MH) algorithm with Hamiltonian Dynamics, such that new proposed values can be further apart than they would be under a random walk exploration scheme, but have a high acceptance probability nevertheless. HMC is then more efficient than MH in the sense that it takes fewer algorithm iterations in order to achieve chain convergence. This is specially relevant in our stochastic volatility context because Bayesian statistics treats latent variables as it treats unknown parameters, what can be computationally expensive. We define daily returns as the first difference of the log-price. Employing natural returns here leads to serious convergence problems in the sampling algorithm and unreliable results.

Preliminary estimation attempts proved parameter λ difficult to identify from data. We choose to set $\lambda = 1$. This affect the strict local martingale criterion of Theorem 1 only if $p = 3/2$. This is not unprecedented. Baldeaux, Ignatieva and Platen (2018) also attempt to empirically identify strict local martingales by estimating restricted versions of model (4.8), all with $\lambda = 1$.

Table 4.3 reports posterior descriptive statistics for Bitcoin. We estimate model (4.8) for each subsample, and for the full sample. In Table 4.4 we show estimates for the four other “traditional” financial assets: the S&P 500 index, the EUR-USD exchange rate, dollar gold prices, and Brent oil prices. Remember we are mostly interested in comparing posterior estimates with the criteria derived by Andersen and Piterbarg (2007), in order to judge whether there are evidences of strict local martingale bubbles in those assets.

Parameter κ gives the speed of mean reversion of the autorregressive process driving volatility V_t . From Table 4.3 we see that for all Bitcoin subsamples, parameter κ has posterior mean between 0.50 and 0.66, but 95% credibility intervals are quite broad, specially for the first and third subsamples. Unconditional mean volatility level is given by parameter θ . Confirming what we see from standard deviations in Table 4.1, θ is estimated largest for the first subsample (0.0715), and smallest for the second subsample (0.0089). Parameter p is the exponent of volatility level V_t , which scales the Brownian innovation dW_t^V in model (4.8). According to the criteria of Andersen and Piterbarg (2007), a value of p between 0.5 and 1.5 characterizes S as a strict local martingale (remember if $p = 0.5$ then model (4.8) reduces to the well known Heston (1993) model). Posterior estimates of p fall wholly in the strict local martingale range, for all Bitcoin subsamples.

Evidence from Bayesian estimation of model (4.8) points to Bitcoin prices behaving like strict local martingales during our first subsample: the 95% credibility interval of parameter p falls in the $(0.5, 1.5)$ range, and correlation ρ is estimated positive with high probability. Similar to the results obtained in the non-parametric estimation of volatility

⁵ See Betancourt (2017) for a thorough explanation of the algorithm.

functions, we did not find robust evidence for the presence of bubbles in the other subsamples of the Bitcoin series, as well as no evidence of bubbles in the traditional assets used as reference.

It is important to note that the stochastic volatility model used represents a good fit for the behavior observed in the Bitcoin series. Figure 5 shows the model fit (estimated for the whole sample) for conditional volatility, compared to Bitcoin's absolute returns series, a proxy for true process volatility. We can observe that the model adequately captures shifts in conditional volatility observed in the Bitcoin series.

Table 4.3 – Posterior Distributions of Andersen and Piterbarg (2007) model - Bitcoin

		Mean	Std	q2.5%	Median	q97.5%
First subsample 2013-01-01:2014-03-31	κ	0.5020	0.2298	0.0634	0.5045	0.9285
	θ	0.0715	0.0562	0.0081	0.0603	0.2082
	ϵ	14.9161	1.6122	11.8690	14.8706	18.1869
	ρ	0.2709	0.1829	-0.0859	0.2643	0.6291
	p	0.7618	0.0406	0.6830	0.7618	0.8422
Second Subsample 2014-04-01:2016-12-31	κ	0.6596	0.2080	0.1982	0.6825	0.9772
	θ	0.0089	0.0093	0.0003	0.0066	0.0304
	ϵ	14.3569	1.5782	11.3895	14.3161	17.5359
	ρ	0.1131	0.1845	-0.2520	0.1144	0.4669
	p	0.7503	0.0338	0.6835	0.7505	0.8159
Third Subsample 2017-01-01:2018-09-01	κ	0.5483	0.2265	0.0899	0.5573	0.9485
	θ	0.0345	0.0330	0.0013	0.0279	0.1064
	ϵ	10.6551	1.6122	7.5858	10.6211	13.9126
	ρ	0.0234	0.2350	-0.4482	0.0237	0.4753
	p	0.8689	0.0458	0.7820	0.8682	0.9608
Full Sample 2013-01-01:2018-09-01	κ	0.6537	0.2162	0.1686	0.6803	0.9770
	θ	0.0040	0.0036	0.0001	0.0028	0.0111
	ϵ	10.6557	1.5564	7.7779	10.6001	13.8711
	ρ	-0.0289	0.1738	-0.3718	0.0128	0.2356
	p	0.8179	0.0331	0.7527	0.8181	0.8825

Note: table reports descriptive statistics of our estimation of Andersen and Piterbarg (2007) model for Bitcoin. First column displays posterior means, second column standard errors, and columns three through display selected quantiles. Estimation was carried out by Hamiltonian Monte Carlo, with one chain of 100000 replications.

Table 4.4 – Posterior Distributions of Andersen and Piterbarg (2007) model - S&P 500, EUR-USD, Gold, and Oil

		Mean	Std	q2.5%	Median	q97.5%
S&P 500	κ	0.6537	0.2162	0.1686	0.6803	0.9770
	θ	0.0040	0.0036	0.0001	0.0028	0.0111
	ϵ	10.6557	1.5564	7.7779	10.6001	13.8711
	ρ	-0.0289	0.1738	-0.3718	0.0128	0.2356
	p	0.8179	0.0331	0.7527	0.8181	0.8825
EUR-USD	κ	0.5941	0.2065	0.2065	0.6030	0.9548
	θ	0.0032	0.0017	0.0017	0.0032	0.0058
	ϵ	6.2593	1.3470	1.3470	6.1781	9.0996
	ρ	0.0224	0.0442	-0.0553	0.0174	0.1218
	p	0.9391	0.0392	0.8603	0.9396	1.0146
Gold	κ	0.6237	0.2020	0.1969	0.6362	0.9647
	θ	0.0051	0.0027	0.0005	0.0052	0.0096
	ϵ	7.1386	1.4592	4.5268	7.0577	10.2148
	ρ	0.0305	0.0726	-0.0941	0.0206	0.2033
	p	0.9406	0.0408	0.8599	0.9411	1.0194
Oil	κ	0.5848	0.2063	0.1559	0.5928	0.9510
	θ	0.0101	0.0060	0.0012	0.0099	0.0208
	ϵ	7.5960	1.3407	5.1814	7.5270	10.4190
	ρ	0.0101	0.0692	-0.1275	0.0083	0.1543
	p	0.9350	0.8550	0.9353	0.9353	1.0127

Note: table reports descriptive statistics of our estimation of Andersen and Piterbarg (2007) model for S&P500, EUR-USD, Gold, and Oil for the sample period 2013-01 until 2018-09. First column displays posterior means, second column standard errors, and columns three through display selected quantiles.

Estimation was carried out by Hamiltonian Monte Carlo, with one chain of 100000 replications.

Figure 4.5 – Fitted Volatility - Andersen and Piterbarg (2007) model

Note: figure plots daily Bitcoin returns and mean posterior conditional volatility (for the model estimated with the full sample). Estimation was carried out by the HMC algorithm with 100000 iterations.

4.4 Concluding Remarks

In this paper we explored the Bitcoin bubble narrative through the strict local martingale theory of financial bubbles (e.g. Jarrow, Protter and Shimbo (2007), Jarrow, Protter and Shimbo (2010), Jarrow, Kchia and Protter (2011b)). A certain financial asset displays bubble pricing if its discounted price is not a “true” martingale under the equivalent neutral probability measure, but a strict local martingale.

Condition (4.7) states an asset price is a strict local martingale if its volatility increases with level faster than linearly. We estimated the volatility function of Bitcoin daily and high-frequency prices for three subsamples, and of other selected financial assets (S&P500, EUR-USD, gold, and Oil) using the nonparametric estimator of Florens-Zmirou (1993). Comparing volatility function estimates between Bitcoin and traditional financial assets we see Bitcoin is much more volatile and displays distinct level-variance dynamics.

Andersen and Piterbarg (2007) show the parameter space of the stochastic volatility model (4.8) can be divided in two disjoint sets, one under which the asset follows a martingale, and another under which the asset follows a strict local martingale. We employed a Hamiltonian Monte Carlo simulation scheme to estimate model (4.8) for each subsample of Bitcoin daily returns, and for those other four selected financial assets. Results point to a bubble in Bitcoin during our first subsample, which goes from January 2013 to April 2014.

5 Foreign Exchange Expectation Errors and Filtration Enlargements

Pedro Chaim
Márcio P. Laurini

Abstract

Extrapolations of future markets forward rates are a better predictor of the 30-days ahead BRL-USD exchange rate than forecasts from the Central Bank Focus survey of Brazilian market participants. This is puzzling because market participants observe forward rates as they submit predictions, and thus these agents perform biased forecasts even though they have access to a set of unbiased forecasts, consistent with a martingale process for the exchange rate. Dandapani and Protter (2016) describe a mechanism through which new information enlarges the information set (a filtration), changing the underlying measure and inducing a drift into the martingale process, turning the process into a strict local martingale and generating a forecast bias. We argue this mechanism can explain our rational conundrum. Empirical results suggest that Focus survey forecasts indeed display characteristics of a strict local martingale, while spot exchange rates and forward rates are consistent with a martingale process.

Keywords: Strict Local Martingales; Filtration Enlargement; Foreign Exchange Markets; Expectation Errors.

5.1 Introduction

Imagine a rational agent that observes some information set (a filtration $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$) and is interested in forecasting values of some random variable. As new information becomes available (in the form of a random variable L , such that we have a new, enlarged filtration $\mathbb{G} = (\mathcal{G}_t)_{t \geq 0}$ where $\mathcal{G}_t = \bigcup_{\epsilon > 0} (\mathcal{F}_{t+\epsilon} \vee \sigma(L))$), one would expect correction of systematic errors and/or some reduction of prediction variance. Through the information accumulation process useless information is discarded, and only information that yields true prediction gains is incorporated — one thus expects forecasts based on the enlarged information set $\mathbb{G}$ to be at least as good as forecasts based on the smaller information set $\mathbb{F}$. In this work we show an example where a larger set of information leads to a bias in the forecasting process, which seems in contradiction with the mechanism described above.

We compare survey-based forecasts from the Market Expectations System of the Central Bank of Brazil (henceforth Focus) with future markets (B3 futures exchange markets, former BMF&BOVESPA) 30-days forward rates; and find that future markets rates are unbiased predictor of future spot rates, but Focus survey forecasts are biased predictors through a substantial portion the observed sample. This is counter-intuitive because forward rates are obviously within the information set of keen market participants. No arbitrage theory asserts that exchange rates should be well characterized by a martingale process under the equivalent risk neutral measure P — i.e., the best possible predictor of future exchange rates is the present value of forward rates. Results in Table 5.2 suggest it is indeed the case in Brazilian markets: it is not possible to reject the martingale hypothesis for either the spot exchange rate, or future market forward rates. Yet Focus survey forecasts seem inconsistent with a martingale process. One could say that markets participants are employing a prediction scheme systematically incompatible with the observed stochastic process of interest. How to explain such phenomenon?

We argue this puzzling phenomenon can be interpreted through strict local martingale theory — specifically through the mechanism detailed by Dandapani and Protter (2016). Dandapani and Protter (2016) explain how enlarging filtration F to G changes the underlying measure (say from P to Q), which can induce a stochastic drift in the volatility of S , turning S from an $\mathbb{F}$ martingale to a $\mathbb{G}$ strict local martingale.

The idea of smaller and larger information sets is natural to problems in finance. The filtration enlargements problem is studied in Jeulin and Yor (1978), Jeulin (1978) and Jeulin and Yor (1985), and more recent developments in Aksamit and Jeanblanc (2017). As an example of application in finance Corcuera and Vadivia (2012) show how filtration enlargements can be applied to model default risk and insider trading problems. Strict local martingales are at the core of a blooming real-time bubble detection literature. For example, Jarrow, Kchia and Protter (2011b) apply a nonparametric methodology (which we borrow in Section 5.4) to show that Linked-In stock exhibited price bubble behavior immediately after the firms IPO, and Chaim and Laurini (2019a) identify Bitcoin bubbles. Recent developments include, for example, Dandapani and Protter (2019) who provide sufficient conditions for the multidimensional case, and Piironen et al. (2018) who introduce a bubble detection indicator based on option prices.

Using two different identification techniques we explore the strict local martingale hypothesis. In Section 5.4 we employ a nonparametric kernel approach based on Jarrow, Kchia and Protter (2011a) to study how the volatility of our series of interest behave as price varies, and access the convergence of condition (5.3). We then turn to an alternative, parametric strict local martingale test in Section 5.5.

To our knowledge this is the first work to relate systematic prediction errors and

the filtration enlargement mechanism. While the nonparametric identification procedure in Section 5.4 draws heavily from Jarrow, Kchia and Protter (2011a), formal univariate estimation of the stochastic volatility model in Section 5.5 is itself novel. Standard practice in financial industry is to calibrate parameter values from option price information (see for example Mikhailov and Nögel (2004)), and classical estimates (like by Ait-Sahalia and Kimmel (2007)) employ volatility proxies to circumvent the latent variable problem — we tackle the latent variable issue by employing Bayesian simulation techniques to estimate model 5.12, so that it is possible to make inference regarding model parameters from a single time series.

The remainder of this paper is structured as follows. Section 5.2 presents some key definitions and a discussion about the filtration enlargement mechanism. Empirical evidence of our main points are found in Section 5.3. Nonparametric strict local martingale identification is presented in Section 5.4. Section 5.5 details our parametric estimation. We discuss results in Section 5.6, and Section 5.7 concludes.

5.2 Filtration Enlargement and Strict Local Martingales

5.2.1 Strict Local Martingales

A local martingale is a stochastic process, $M = (M_t)_{t \geq 0}$, for which there exists an increasing sequence of stopping times $\tau_n \rightarrow \infty$ such that the stopped process $M_{\min\{\tau, \tau_n\}}$ is a martingale with probability one. The key insight is that if M is a local martingale, then it is either a true martingale or a strict local martingale (SLM).

The risky asset price, $S = (S_t)_{t \geq 0}$, is assumed non-negative and driven by a standard stochastic differential equation

$$dS_t = \sigma(S_t)dW_t + \mu(S_t)dt. \quad (5.1)$$

No arbitrage in the sense of “No Free Lunch with Vanishing Risk” (Delbaen and Schachermayer (2006)) implies that there exists a risk neutral measure under which the SDE in (5.1) simplifies to

$$S_t = S_0 + \int_0^t \sigma(S_s)dW_s. \quad (5.2)$$

A very important property (to be further exploited) is that S is a strict local martingale if and only if

$$\int_\alpha^\infty \frac{x}{\sigma(x)^2} dx < \infty \quad (5.3)$$

for all $\alpha > 0$. Jarrow (2012) notes that condition (5.3) states that an asset is a strict local if and only if its volatility, as function of its price, increases faster than linearly.

Jarrow, Kchia and Protter (2011a) discuss the intuition behind the distinction between a non-negative martingale and a strict local martingale, $S \geq 0$, follows from the fact that S is always a supermartingale and is a true martingale if and only if it has constant expectation. Then if S is a strict local martingale its expectation decreases over time. The typical behavior of a SLM process is to shoot up to high values, to then drop down and stay low.

5.2.2 Filtration Enlargement

Dandapani and Protter (2016) detail a mechanism through which new information introduced into the observed information set (the initial filtration $\mathbb{F}$) induce a change in the measure, from P under $\mathbb{F}$ to Q under the enlarged filtration $\mathbb{G}$, such that a $\mathbb{F}$ martingale process S can become a $\mathbb{G}$ strict local martingale for a random time interval.

To understand how the filtration enlargement mechanism connects strict local martingale bubbles and systematic foreign exchange forecast errors, imagine some agent One wants to predict future values of a risky asset $S = (S_t)_{t \geq 0}$ (for example the 30-days spot BRL-USD exchange rate). Agent One observes an information set (a filtration) $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$, such that we have a complete filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, P)$. Suppose S is a $\mathbb{F}$ true martingale under measure P .

Now consider some other agent Two that also aims to forecast S . At time $t = 0$ only agent Two observes a random variable L , that enlarges her information set, from $\mathbb{F}$ to $\mathbb{G} = (\mathcal{G}_t)_{t \geq 0}$ — i.e., $L \in \mathbb{G}$ but $L \notin \mathbb{F}$. Dandapani and Protter (2016) show that it is possible for, due to such change, process S to be a $\mathbb{G}$ strict local martingale. The mechanism that allows the construction of a strict local martingale is a class of stochastic volatility model, detailed below.

Consider the Heston type model studied by Lions and Musiela (2007)

$$dS_t = S_t v_t dB_t \tag{5.4}$$

$$dv_t = \sigma(v_t) dW_t + b(v_t) dt \tag{5.5}$$

where dB_t and dW_t are two correlated Brownian motions with correlation parameter ρ . Denote the new, enlarged filtration $\mathbb{G}$ as

$$\mathcal{G}_t = \bigcap_{\epsilon > 0} (\mathcal{F}_{t+\epsilon} \vee \sigma(L)).$$

Let η be the distribution of L , and $Q_t(w, dw)$ be the regular conditional distribution of L given $\mathcal{F}_t$. There exists an $\mathbb{F}$ martingale $q^x(t, w)\eta(d_x)$ which is a version of $Q_t(w, dx)$. If S is an $\mathbb{F}$ martingale, then there exists an F predictable process $k^x(t, 2)$, such that

$$[q^x, S] = (k^x q_-^x) \cdot [S, S] \tag{5.6}$$

Jacod (1985) shows that the following process $\hat{S}_t$ is a $\mathbb{G}$ local martingale

$$\hat{S}_t = S_t - \int_0^t k_d^L d[S, S]_s. \quad (5.7)$$

Enlarging filtration $\mathbb{F}$ to $\mathbb{G}$ changes the semimartingale decomposition of both S and v . Girsanov's theorem allows us to restore the local martingale properties of S by changing the equivalent probability measure to Q .

Under $(Q, \mathcal{G})$, S can be written as

$$\begin{aligned} S_t = & \int_0^t (S_s v_s) dB_s - \int_0^t k_s^L (S_s v_s)^2 ds \\ & - \int_0^t ((S_s v_s) H_s + \rho J_s) ds + \int_0^t k_s^L (S_s v_s)^2 ds + \int_0^t ((S_s v_s) H_s + \rho J_s) ds \end{aligned} \quad (5.8)$$

to ensure that S is a Q local martingale we choose

$$k_t^L (S_t v_t)^2 = -(S_t v_t) H_t - \rho J_t. \quad (5.9)$$

The drift component $b(v_t)$ in volatility equation (5.5) also changes

$$\hat{b}(v_t) = b(v_t) + k_t^L \sigma^2(v_t) + (\rho H_t + J_t) \sigma(v_t). \quad (5.10)$$

This is the change in the drift of the volatility equation that enables the change from a martingale to a strict local martingale.

We can now state Theorem 3 of Dandapani and Protter (2016) (p.p. 10), which relates S local martingale under $(\mathbb{F}, P)$ to S strict local martingale under $(\mathbb{G}, Q)$.

Theorem 1. *Dandapani and Protter (2016). Assume that k, H and J have right continuous paths almost surely, and $Q(w : k_0^L > 0) > 0$. Suppose functions $\sigma(\cdot)$ and $b(\cdot)$ are $\mathcal{C}^1$, and the following two conditions hold:*

$$\limsup_{x \rightarrow +\infty} \frac{\rho x \sigma(x) + b(x)}{x} < \infty \quad (i)$$

$$\liminf_{x \rightarrow +\infty} \left(\rho x \sigma(x) + b(x) \right) + \min(\varepsilon^{(1)}, \varepsilon^{(2)}) \sigma^2(x) - \max(\varepsilon^{(1)}, \varepsilon^{(2)}) \sigma(x) \Big) \phi^{-1}(x) \geq 0 \quad (ii)$$

where $\phi(x)$ is an increasing, positive, smooth function satisfying

$$\int_\alpha^\infty \frac{1}{\phi(x)} dx < \infty.$$

Let S be the unique strong solution of the SDE

$$\begin{aligned} dS_t &= S_t v_t dB_t \\ dv_t &= \sigma(v_t) dW_t + b(v_t) dt \end{aligned}$$

on $(P, \mathbb{F})$, where B_t and W_t are Brownian motions with correlation $\rho > 0$. Then process S is also the solution of

$$\begin{aligned} dS_t &= S_t v_t dB_t \\ dv_t &= \sigma(v_t) dW_t + b(v_t) dt + k_t^L \sigma^2(v_t) dt + (\rho H_t + J_t) \sigma(v_t) dt. \end{aligned}$$

Proof: See Dandapani and Protter (2016).

Therefore if functions $\sigma(\cdot)$ and $b(\cdot)$ respect conditions (i) and (ii), then under the enlarged filtration $\mathbb{G}$ the drift component $b(\cdot)$ changes to $\hat{b}(\cdot)$ (as in equation (5.10)), which turns S from an $\mathbb{F}$ (true) martingale into a $\mathbb{G}$ strict local martingale.

The presence of stochastic volatility in financial series is an important stylized fact in the field of finance. In the following sections we show that the characterization of a strict local martingale through a stochastic volatility model is consistent with the behavior observed in the forecasts for the foreign exchange market studied in this article, and thus can be a valid explanation for the observed forecast bias.

5.3 Empirical Evidence

Our sample consists of (30-days ahead) exchange rate predictions of Brazilian market participants (the Focus survey); future markets closing forward rates traded at B3 (former Bovespa BM&F); and the official Brazilian exchange rate (PTAX). Sample range and frequency are limited by Focus data such that 179 monthly observations, from 2001/11 to 2016/09, are available. The data used in this work are available on the websites of Central Bank do Brazil and B3.

Focus forecasts regard the last week day of the corresponding month, while future markets forward rates are for the first week day of the same month. We thus obtain the “predicted forward rate at the last day of the month” by using the CDI (Brazilian Interbank Deposit Certificate) rate. We also consider two alternative formulations — first using the forward rate the day before Focus release, and then using the forward two months maturity. Results are similar to the baseline formulation and therefore not reported. Trajectories are plotted in Figure 5.1, and descriptive statistics are displayed in Table 5.1.

The Focus Bulletin is a survey released every Monday by the Central Bank of Brazil. The report, a daily survey of forecasts of about 120 banks, funds and other institutions summarizes the market expectations for the Brazilian economy. The Market Expectations System, which collects information reported in the Focus Bulletin, consists of a web interface where agents report their expectations for the macroeconomic variables of interest. The system considers the last information reported in the past 30 days by

Figure 5.1 – Spot rate, Focus and future market forecasts

Note: This figure plots observed values of BRL-USD (PTAX) spot exchange rates, BMF&BOVESPA CDI-discounted forward rates, and Focus forecasts.

market participants. If an agent does not revise his or her expectations in the last thirty days, this information is not considered in statistic calculations. An important incentive of this system is the monthly release of best performing institutions, indicated in the top five predictors for each macroeconomic variable monitored in the system (see Mariani and Laurini (2016) for a thorough description of the Focus survey).

Table 5.1 – Descriptive statistics

	Mean	SD	Max	Min	MFE
Spot	2.397	0.630	4.117	1.564	0.000
Focus	2.393	0.625	4.080	1.570	-0.004
Future	2.414	0.636	4.079	1.572	0.017

Note: This table reports descriptive statistics for BRL-USD (PTAX) spot exchange rates, BMF&BOVESPA CDI-discounted forward rates, and Focus forecasts. Mean forecast error (MFE) is with respect to spot exchange rates.

We now verify whether processes describing forward rates, Focus forecasts, and the spot exchange rate are consistent with a martingale process — as expected by the first theorem of asset pricing. For such we employ test based on Cramer-von Mises (CvM) and Kolmogorov-Smirnov (KS) statistics proposed by Domínguez and Lobato (2003), augmented with bootstrap finite sample correction of Charles, Darné and Kim (2011). Test statistics and equivalent p-values are reported in Table 5.2. Both tests fail to reject the martingale null hypothesis for the spot exchange rate and future markets forward rates, but reject it when considering Focus forecasts. The same tests are applied to prediction errors (Future-spot and Focus-spot). We find that the difference martingale hypothesis is

Figure 5.2 – Rolling sample bias test

Note: This figure plots rolling sample bias test for Focus forecasts. P-values based on a rolling sample estimation.

rejected for Focus prediction errors, and not rejected for future market prediction errors. This result indicates that there is a systematic bias component in the forecasts reported in the Focus system, which is inconsistent with the behavior of spot and future exchange rates and indicating that Focus forecasts are sub-optimal.

Table 5.2 – Martingale tests

	CvM	CvM p-val	KS	KS p-val
Focus	2.569	0.000	2.282	0.000
Future	0.042	0.948	0.553	0.897
Spot	0.070	0.810	0.731	0.616
Δ Focus	2.730	0.000	2.150	0.000
Δ Future	0.026	0.990	0.549	0.910

Note: This table reports martingale tests of BRL-USD (PTAX) spot exchange rates, B3 (BMF&BOVESPA) CDI-discounted forward rates, and Focus survey forecasts. Cramer von Mises (CvM) and Kolmogorov-Smirnov statistics calculation were based in 2000 bootstrap replications.

Figure 5.2 plots p-values for a rolling-sample forecast bias test. We regress 30-days ahead exchange rates against Focus forecasts and a constant, i.e. $y_t = \alpha + \beta x_t + \varepsilon_t$, and jointly test if $\alpha = 0$ and $\beta = 1$. From 2001/11 to 2008/07 it is possible to reject the null $\alpha = 0$ and $\beta = 1$ hypothesis at a 10% significance level, again supporting the existence of a relevant bias in the forecasts of market analysts reported in the Focus system.

Table 5.3 displays an illustrative prediction performance comparison between Focus survey forecasts and future market rates. We repeat the procedure for the "biased" and "unbiased" Focus subsamples (2001/11/30-2008/07/31 and 2008/09/29-2016/09/30). Forward rates are generally a better predictor the 30-days ahead spot interest rate. When

considering the whole sample or any subsample forward rates fare better when considering RMSE or MAE, while Focus mean prediction errors are smaller in absolute terms.

Table 5.3 – Forecast comparison

	ME	RMSE	MAE
Focus	-0.004	0.159	0.102
Future	0.017	0.146	0.096
Focus first subsample	0.011	0.182	0.110
Future first subsample	0.035	0.167	0.105
Focus second subsample	-0.016	0.136	0.095
Future second subsample	0.002	0.126	0.088

Note: This table reports BRL-USD (PTAX) spot exchange rates prediction performance comparison statistics of BMF&BOVESPA CDI-discounted forward rates, and Focus forecasts. We split the observed sample in two sub-samples, 2001/08 to 2008/07, and from 2008/08 forwards. Column one displays mean error, second column root mean squared error, third column mean absolute error.

For each subsample we test the null hypothesis of no difference in forecast accuracy by using the Diebold and Mariano (2002) and Clark and West (2007) tests. From p-values in Table 5.4 we can see the null of equal forecast performance is rejected at 5% by both tests for all subsamples. Indeed, Clark and West (2007) tests reject the null at the 1% level for all subsamples; and p-values of Diebold and Mariano (2002) are always smaller than 0.03. These results reinforce forward rates are indeed better predictors of 30-days ahead spot rates than Focus survey forecasts.

In the following sections we show that this bias component can be explained by the presence of a strict local martingale, using the framework proposed in Dandapani and Protter (2016). For this we use non-parametric and parametric tests for the presence of an SLM, verifying the existence conditions in the volatility structure of the price process.

Table 5.4 – Forecast tests

	DM	CW
Full Sample	0.0154	<1e-3
First Subsample	0.0294	<1e-3
Second Subsample	<1e-3	<1e-3

Note: This table displays p-values of Diebold and Mariano (2002) (DM) and Clark and West (2007) (CW) tests. The DM test has the null hypothesis of equality of predictive performance, against the alternative of better predictive performance for the first model (Future market forecasts). Tests employ quadratic loss function.

Figure 5.3 – Variance functions, whole sample

Note: This figure plots volatility function $\sigma(S_t)$ (top panel) and relation (5.3) (bottom panel) of spot exchange rates, discounted forward rates, and Focus forecasts.

5.4 Nonparametric Estimation

Recall from our discussion in Section 5.2 that a stochastic process is a strict local martingale if its volatility increases fast enough with higher level, such that

$$\int_{\alpha}^{\infty} \frac{x}{\sigma^2(x)} dx < \infty \quad (5.3)$$

for all $\alpha > 0$. We try to empirically characterize the volatility term $\sigma(x)^2$ in order to study the convergence properties of the integral above. In order to estimate $\sigma^2(S_t)$ and relation (5.3), we employ the kernel-based estimator of Florens-Zmirou (1993) and Jiang and Knight (1997) for the diffusion component:

$$\hat{\sigma}^2(x) = \frac{\sum_{i=0}^{n-1} K\left(\frac{x-X_i}{h_n}\right) (X_{i+1} - x)^2}{\Delta_n \sum_{i=0}^{n-1} K\left(\frac{x-X_i}{h_n}\right)} \quad (5.11)$$

where $K(\cdot)$ is a kernel function with bandwidth h_n . In this work we use a Gaussian kernel and generalized cross-validation to estimate the optimal bandwidth parameter.

Figure 3 plots estimated volatility functions (top panel) and relation $x/\sigma^2(x)$ (bottom panel), for our three processes of interest. One readily notices that Focus forecasts are less volatile than spot exchange rates or future market predictions for every observed price realization, and that volatility functions of both spot rates and future markets stop increasing somewhere past the $x = 3.75$ mark. The relation $x/\sigma^2(x)$ is stable (one might even say increasing) at higher price ranges for spot and future processes, which alludes to a violation of condition (5.3) (i.e., the integral is infinite) — indicating that Spot rates and future market predictions are true martingales.

Figure 5.4 – Vertically rescaled variance functions, whole sample

Note: This figure plots vertically rescaled variance function $\sigma^2(S_t)$ (top panel) and relation (5.3) (bottom panel) of spot exchange rates, discounted forward rates, and Focus forecasts.

The curvature of the volatility function of Focus forecasts decreases slower, such that the function is always increasing for observed prices. This reduces the term $x/\sigma^2(x)$ and indicates convergence of the integral in (5.3). We argue systematic Focus prediction errors are associated with incompatibility between the stochastic process that drives forecasts and exchange rates. Figure 4.3 plots $\sigma^2(x)$ and $x/\sigma^2(x)$ for the first subsample period. Note the absence of inflection at the end of Focus' variance function, and how sharply $x/\sigma^2(x)$ decreases. This can be taken as evidence that Focus forecasts behave like strict local martingales from 2001/11 to 2008/08. Conversely, when considered the second subsample, as in Figure 4.4, it becomes hard to distinguish between Focus and spot processes — both are consistent with what is to expect from true martingales.

This nonparametric analysis has some important limitations. First, the analysis is visual and presents no formal test for the SLM hypothesis. The second point is that the SLM behavior refers to extrapolation for the whole set of unobserved prices, whereas we only analyze the variance function for observed sample values. To circumvent this limitation, we present a formal parametric test, detailed in the next section.

5.5 Parametric Estimation

Consider the stochastic volatility model studied by Andersen and Piterbarg (2007)

$$\begin{aligned} dS_t &= \lambda S_t \sqrt{V_t} dW_t^1 \\ dV_t &= \kappa(\theta - V_t)dt + \epsilon V_t^p dW_t^2 \end{aligned} \tag{5.12}$$

Figure 5.5 – Vertically rescaled variance functions, first subsample

Note: This figure plots vertically rescaled variance function $\sigma^2(S_t)$ (top panel) and relation (5.3) (bottom panel) of spot exchange rates and Focus forecasts for the first subsample period.

Figure 5.6 – Vertically rescaled variance functions, second subsample

Note: This figure plots vertically rescaled variance function $\sigma^2(S_t)$ (top panel) and relation (5.3) (bottom panel) of spot exchange rates and Focus forecasts for the second subsample period.

where $\lambda, \kappa, \theta > 0$ and (W_t^1, W_t^2) is a standard two dimensional Brownian motion with correlation coefficient ρ . Note that if $p = 1/2$ the model in (5.12) reduces to the well-known Heston (1993) model. This model is equivalent to the one discussed in Section 5.2, but also allows for testable implication of the Martingale hypothesis. Results of Andersen and Piterbarg (2007) and Protter (2013) describe how the parameter space of model (5.12) can be divided in two disjoint sets, such that a specific parameter range characterizes stochastic process S as a strict local martingale.

Theorem 2. *Andersen and Piterbarg (2007). For the model (5.12): if $\rho \leq 0$, S is a true martingale; if $\rho > 0$ and $p \leq 1/2$ or $p > 3/2$, S is a true martingale; if $\rho > 0$ and $1/2 < p < 3/2$, then S is a strict local martingale. For the case $p = 3/2$ S is a true martingale if $\rho \leq 1/2\epsilon\lambda^{-1}$, and S is a strict local martingale if $\rho > 1/2\epsilon\lambda^{-1}$. Proof: see Protter (2013)*

Equipped with this theorem we employ Bayesian techniques to estimate the stochastic volatility model of Andersen and Piterbarg (2007) to investigate whether Focus forecasts are compatible with a true martingale or a strict local martingale process.

Estimation was carried out in the Stan software¹. Stan employs a Hamiltonian Monte Carlo (HMC) algorithm rather than the traditional Metropolis Hastings (MH). HMC augments the usual random walk posterior sampling scheme with Hamiltonian dynamics, such that new proposed values can be further apart than they would be under a random walk exploration scheme, but have a high acceptance probability nevertheless. HMC is then more efficient than MH in the sense that it takes fewer algorithm iterations in order to achieve chain convergence. This is especially relevant in our stochastic volatility context because Bayesian statistics treats latent variables as it treats unknown parameters, what can be computationally expensive.

We elected prior distributions given restrictions on parameter values in model (5.12). A Uniform distribution between 0 and 1.5 was specified for λ and ϵ ; κ was assumed Normal, with mean 0.9 and standard deviation 0.5; θ was assumed Normal, with mean 0.01 and standard deviation 0.05; p was also assumed Normal, with mean 0.5 and standard deviation 0.05; correlation ρ was specified as $2\rho^* - 1$, $\rho^* \sim \text{Beta}(5, 2.5)$.

Table 5.5 reports estimation results, posterior densities are displayed in Figure 5.7. Notice correlation coefficient ρ is estimated positive inside the 95% credible interval — this is desirable since if $\rho \leq 0$, S_t is a true martingale for all values of p , and further analysis would be unnecessary.

The posterior mean of parameter p is estimated around 0.365, which according to Theorem 2 characterizes Focus forecasts as a true martingale process. However, the

¹ With help of the R-package “rstan”

Table 5.5 – Estimation results, unrestricted model

	Mean	Std	.025q	.5q	.975q
λ	0.630	0.378	0.137	0.552	1.416
κ	0.748	0.194	0.280	0.792	0.990
θ	0.168	0.119	0.048	0.123	0.468
ϵ	0.826	0.348	0.273	0.799	1.452
ρ	0.226	0.180	0.026	0.174	0.667
p	0.365	0.204	0.028	0.364	0.743

Note: This table report results for the unrestricted estimation of model (5.12). First column displays posterior means, second column standard errors, and columns three through five display selected quantiles. Estimation was carried out by Hamiltonian Monte Carlo with 100000 replications.

Figure 5.7 – Posterior Densities, unrestricted model

Note: This figure plots posterior distributions of the unrestricted model (5.12), estimated via Stan with 100000 replications.

posterior distribution of p is quite disperse and imply substantial uncertainty regarding parameter value. Also, prior and posterior distributions of parameters λ and ϵ are very similar, thus it is questionable whether those parameters are identifiable from data. We recognize that available Focus data is of monthly frequency and sample size is limited, which makes estimating model (5.12) an ambitious enterprise. In face of this identification difficulty we decide to fix $\lambda = 1$ and focus attention on estimates of p and ρ , which are at the core of our parametric approach. This restriction is also used in Baldeaux, Ignatieva and Platen (2018) and Chaim and Laurini (2019a), which also attempt to verify the presence of strict local martingales.

One can see from Table 5.6 that the restricted model indeed yields more precise estimates. Comparing the second column of Table 5.5 and Table 5.6 one can see that posterior standard errors of θ , ρ , and p are substantially smaller; while κ estimates are essentially unchanged. Figure 5.8 displays bell-shaped, less disperse, posterior distributions

Figure 5.8 – Posterior Densities, restricted model

Note: This figure plots posterior distributions of the restricted $\lambda = 1$ model (5.12), estimated via Stan with 100000 replications.

of κ , ρ , and p .

Table 5.6 – Estimation results, restricted model

	Mean	Std	.025q	.5q	.975q
κ	0.743	0.199	0.262	0.788	0.991
θ	0.095	0.037	0.048	0.088	0.186
ϵ	0.789	0.291	0.340	0.747	1.408
ρ	0.194	0.093	0.062	0.178	0.415
p	0.507	0.046	0.413	0.507	0.597

Note: This table report results for the restricted ($\lambda = 1$) estimation of model (5.12). First column displays posterior means, second column standard errors, and columns three through five display selected quantiles. Estimation was carried out by Hamiltonian Monte Carlo with 100000 replications.

Parameter p is now estimated between 0.4 and 0.6 with 95% probability, a much narrower credibility interval. Still, because substantial probability is attributed to the $p \leq 0.5$ range, it is not possible to convincingly assert that S_t is a strict local martingale process from Table 5.6 alone. A more formal test is appropriate. We ask: does imposing that S_t is a strict local martingale worsen fit performance? for such we further restrict model (5.12) such that $p \in [0.5, 1.5]$. Then compare this formulation to the one in Table 5.6 and Figure 5.8, for which $p \in [0.5, \infty)$. We evaluate model performance through the Leave-one-out cross-validation (LOO) and the widely applicable information criterion (WAIC) of Vehtari, Gelman and Gabry (2016) — those are pointwise out-of-sample prediction accuracy tests of a fitted Bayesian model, and are based on the log-likelihood evaluated at the posterior simulation of the parameter values.

Figure 5.9 – Fitted Volatility

Note: This figure plots fitted volatility versus absolute returns for the $\lambda = 1$ models.

Table 5.7 reports WAIC and LOO information criteria for the unrestricted, $p \in [0, \infty)$, and restricted, $p \in [0.5, 1.5]$ models. Imposing SLM does not worsen out of sample forecast behavior, as both WAIC and LOO attribute a smaller information criterion to the restricted model. Figure 5.9 plots fitted versus realized Focus volatility for both models in Table 5.7.

Table 5.7 – Model evaluation

	$p \in [0, \infty)$	$p \in (0.5, 1.5)$
WAIC	-659.8 [28.3]	-663.8 [28.9]
LOO	-613.3 [36.3]	-620.1 [36.4]

Note: This table reports WAIC and LOO information criteria estimates (standard errors in square brackets). First column display results for the p -unrestricted model, and second column displays results for the SLM imposed model.

5.6 Discussion

Although one must recognize methodological choices and statistical uncertainty, we view the results of Sections 5.4 and 5.5 as evidence that Focus survey forecasts are

better characterized by a strict local martingale than a driftless true martingale process, specially through the first subsample from 2000/08 to 2008/07.

The nonparametric approach of empirically evaluating condition (5.3) is fundamentally limited by the range of observed values S_t in a given time interval. We are trying to extrapolate from a compact set of R_+ properties of $\sigma^2(S_t)$ for all of R_+ , and, as discussed in Jarrow, Kchia and Protter (2011b) results will invariably depend on the extension method employed.

Theorem 2 gives us empirically testable objective criteria that characterize a price process S_t as either a true martingale or a strict local martingale. But formally estimating Heston-type models is not standard practice for understandable reasons. Classical likelihood inference has major difficulties when dealing with latent variables, what makes Bayesian simulation methods appealing.

When we compromise and fix $\lambda = 1$, our estimates are still inconclusive with respect to Theorem 2 — because although one can say with reasonable certainty that ρ is positive, the posterior of p attributes significant probability to the true martingale range, what makes a more formal test of the strict local martingale hypothesis warranted.

We draw from developments in Bayesian model selection (Vehtari, Gelman and Gabry (2016)) and devise an experiment in which one imposes that S_t is a strict local martingale (i.e. $\rho > 0$ and $0.5 < p < 1.5$), and evaluate whether this restriction hinders out-of-sample forecast performance. Imposing SLM behavior does not worsen out-of-sample forecast, but it is hard to distinguish between models. In hindsight this result is not surprising because even in the p -unrestricted model more, probability is attributed to the SLM range of $p \in (0.5, 1.5)$ than otherwise.

5.7 Concluding Remarks

We investigate the puzzling poor performance of Brazilian market participants in forecasting the BRL-USD exchange rate. Extrapolations of forward rates are a better predictor of the 30-days ahead spot exchange rates than Focus survey forecasts — which are also biased through a substantial portion of our sample, from August 2002 to July 2007. This result is counter-intuitive since Focus participants have access to the information on the future foreign exchange market, and yet they make biased forecasts.

Dandapani and Protter (2016) describe a mechanism through which the incorporation of new information induces a drift in an otherwise driftless local martingale, turning it into a strict local martingale. We argue this filtration enlargement mechanism could help explain our forecast performance conundrum. Spot and future rates are consistent with martingales, while Focus forecasts have a bias term, equivalent to the presence of a drift.

Another important feature is the presence of stochastic volatility in this process, another important feature in the characterization of a strict local martingale used by Dandapani and Protter (2016) to represent the drift term generated by the process of information acquisition (filtration enlargement).

In order to investigate the strict local martingale hypothesis, we employ two different identification methods. The first one is a nonparametric estimation of variance function $\sigma^2(S_t)$ and rely on condition (5.3) — that asserts a local martingale is a strict local martingale if its volatility increases faster than linearly with level. Focus forecasts are less volatile, in absolute terms, than either spot or future rates, but Focus volatility increases faster for higher price observations. The second draws on results of Andersen and Piterbarg (2007): depending on parameter values of the stochastic volatility model (5.12), the underlying price process is either a true martingale or a strict local martingale. Employing Bayesian simulation techniques we estimate Focus 30-days ahead exchange rate forecasts follow a strict local martingale with higher probability than otherwise.

In this paper we present a new way of interpreting the presence of biased forecasts in a market where there are alternative forecasting methods that are informationally efficient. Using the filtering process proposed by Dandapani and Protter (2016), we show that in the presence of stochastic volatility it is possible that more informed forecasts are worse than predictions based on a set of more restricted information. This result is important for the analysis of learning processes and information acquisition in financial markets.

6 Estimating long memory stochastic volatility models using integrated nested Laplace approximations

Pedro Chaim
Márcio P. Laurini

Abstract

In this work we discuss the use of integrated nested Laplace approximations to perform the estimation of stochastic volatility models with long memory. Log volatility persistence is given by a Fractional Gaussian Noise process, which is approximated as a linear combination of independent first-order autoregressive processes that has a Gaussian Markov Random Field representation. We show that this approach is computationally efficient, and performs comparable to a Markov Chain Monte Carlo estimator. We used this methodology to estimate the volatility dependency patterns for the SP&500 index and the most important cryptocurrencies, analyzing the in-sample fit and the construction of out-of-sample forecasts. We also show that this method can be used for the construction of realized volatility measurements using high frequency data, showing the optimal computational performance of the proposed method.

Keywords: Long Memory; Gaussian Markov Random Field; Laplace Approximations; Volatility Forecasting.

6.1 Introduction

Conditional volatility models are an essential tool in the field of finance econometrics and risk management, and because of this importance there is a long literature related to the development of econometric models and inference methods to capture volatility patterns over time. The main problem is that conditional volatility is a latent process, which adds a considerable degree of difficulty to traditional inference methods, such as maximum likelihood estimators. The literature on econometric models of conditional volatility can be divided into two main classes of models based on the form of treatment for the latent volatility process. The first important class is the class of observation driven models, which considers that the process of latent volatility can be approximated by a structure that depends only on observable processes. In this class of models the main reference is the

GARCH (Generalized Autoregressive Conditional Heteroskedasticity) family of models, introduced by Engle (1982) and generalized by Bollerslev (1986) and later extended in several other directions.

In this structure, latent variance is considered a deterministic function of the past variance itself and the past squared returns or residues. A general framework for observation driven process modeling is the generalized score model class proposed in Creal, Koopman and Lucas (2013), which uses the model's past score as the update mechanism for time-varying model parameters. The use of this observation-driven structure is interesting because considering that the volatility process depends only on observable components in the current period, it avoids the inference problems generated by the presence of latent variables in the model, allowing the direct use of maximum likelihood estimators.

The second class of models is based on the treatment of conditional volatility as a truly stochastic dependent process, the main reference in discrete time being the so-called log-normal model of stochastic volatility introduced by Taylor (1986). In this model the conditional log-variance has its dynamics given by a first-order autoregressive process. This structure is very convenient since this formulation is consistent with the main stylized facts of financial series, such as volatility clusters, heavy tails, and non-Gaussianity of returns, and avoids the artificial assumption that volatility is deterministic.

However, assuming that the process of latent volatility is stochastic adds significant difficulties in the inference procedures for parameters and for the filtering of the conditional volatility itself. When assuming a stochastic structure it is not possible to directly use maximum likelihood estimators, since these require the marginalization of each unobserved volatility, thus necessitating the calculation of a multiple integral with a dimension equal to the sample size, which is not feasible computationally. Due to this difficulty a series of frequentist and Bayesian methods was created to carry out the estimation of stochastic volatility models. The estimation method proposed in Taylor (1986) is based on the method of moments, which does not allow to estimate the latent volatility process itself and only the parameter vector of the model. Among frequentist methods the main reference is the Quasi-Maximum likelihood method based on the decomposition of the forecast error through the Kalman Filter independently proposed by Harvey, Ruiz and Shephard (1992) and Nelson (1988). In this structure the log of the squared returns is composed of a mean component plus an latent autoregressive process that is the log-variance of the returns, formulated as a state-space process that allows the use of the Kalman Filter for the filtering of this latent component. As this estimation is based on a linearization of the model, the estimator is approximated and therefore is a Quasi-Maximum Likelihood estimator. Although this estimator is computationally simple, its performance may be sub-optimal due to the bias introduced by the linearization of the model, as discussed in Andersen and Sørensen (1996). To circumvent this limitation, other alternative frequen-

tist estimation methods use maximum simulated likelihood (Sandmann and Koopman (1998)), efficient method of moments (Andersen, Chung and Sørensen (1999)), empirical likelihood and minimum discrepancy (Laurini and Hotta (2017)), among other possible ways.

Concurrently several Bayesian estimation methods have been proposed for the class of stochastic volatility models. Bayesian estimation is particularly interesting for this problem since latent processes can be treated as additional components of the model, and thus additional parameters to be estimated, which allows the use of traditional methods of Bayesian inference without additional complications. Among the main works proposing the use of Markov Chain Monte Carlo for estimating SV models we have Kim, Shephard and Chib (1998), Chib, Nardari and Shephard (2002) and Asai (2005), and as examples of recent developments Kastner and Frühwirth-Schnatter (2014) and Kastner (2019).

An alternative form of Bayesian estimation of SV models was proposed in Martino et al. (2011), through the use of Integrated Nested Laplace Approximations (INLA), a class of methods proposed in Rue, Martino and Chopin (2009). The INLA methodology allows to perform the Bayesian estimation for parameters and latent factors for a wide class of models that can be represented or approximated by Gaussian Markov Random Fields. This approximation is analytical, and does not require the use of simulation methods, as used in MCMC algorithms, and thus is not affected by chain convergence problems. In addition, these methods are very accurate and generally computationally fast and efficient, since in conjunction with the INLA methodology it is possible to use a series of sparse matrix representations to represent the Gaussian Markov Random Field. The good properties of this method for the estimation of SV models are analyzed in Ehlers and Zevallos (2015), which show that the method is very accurate and has good properties for the calculation of measures derived from the conditional volatility estimated as the Value At Risk.

An important point is that most of the models proposed for stochastic volatility modeling are based on a short memory structure for the dependency process, assuming an mean reversion velocity compatible with an autoregressive process. An alternative possibility is to assume that the process of dependence on volatility is given by a long memory process, also known as a long range dependence process. In this structure the most used form is through an innovation structure based on a discrete version of the fractional Brownian Motion, through a representation based on a fractional Gaussian Noise, as described in Taqqu (2003). In this case the innovations of the dependency process follow a hyperbolic decay process, rather than the standard exponential decay of short memory models. References on the use of long memory structures in stochastic volatility models can be found in Harvey (1998), Breidt, Crato and Lima (1998) and Hurvich and Soulier (2009). The main model used is the Long Memory Stochastic Volatility (LMSV), where

the log variance follows a long memory process.

The use of long memory models for variance may be interesting as there is empirical evidence showing that this structure may be appropriate for some sets of assets and measures of volatility. Christensen and Nielsen (2007), Scharth and Medeiros (2009), Hillebrand and Medeiros (2016), Shackleton, Pong and Taylor (2008) and Fleming and Kirby (2011) discuss the importance of long memory presence in conditional stock market volatility. Another important point is the impact of long memory presence on the calculation of realized volatility measurements using high frequency data, as discussed in Maasoumi and McAleer (2008), McAleer and Medeiros (2008) and Lieberman and Phillips (2008). There is also a recent interest in the use of these models because of the ample evidence of the presence of long memory in the cryptocurrency market, as shown by Mensi, Al-Yahyaee and Kang (2018), Phillip, Chan and Peiris (2019), Bouri et al. (2019) and Chaim and Laurini (2019b) among other works.

The modeling of stochastic volatility with the use of a long memory structure presents an additional complexity in relation to the models that assume an autoregressive structure. The key point is that the long memory structure based on Fractional Brownian Motion/Fractional Gaussian Noise is a non-Markovian and non-semimartingale process, violating assumptions normally assumed of a Markovian structure and Martingale difference innovation processes used in the estimation of these models. As a direct example of the difficulties generated by the introduction of long memory structures, note that a good part of the methods used in the estimation of stochastic volatility models is based on a linear Gaussian state space representation, whose fundamental assumption is the use of a Markovian structure of dependence for latent states. Due to this this difficulty, alternative forms were proposed for the estimation of this class of models.

Harvey (1998) and Breidt, Crato and Lima (1998) use a spectral approximation for the likelihood function; Arteche (2004) a semi-parametric form based on a local Whittle estimator. Different formulations can be used using an approximation of the state space representation through a truncated ARFIMA process, as used in Chan and Palma (1998), Basak, Chan and Palma (2001) and Ferraz and Hotta (2007) in frequentist formulations, and Chan and Petris (2000) for the Bayesian estimation in a Markov Chain Monte Carlo algorithm. Comparative reviews of estimation methods of stochastic volatility models with long memory can be found in Perez and Ruiz (2001) and Crato and Ray (2002).

In this paper we propose an alternative Bayesian estimation for long memory stochastic volatility models using an approximation for a Fractional Gaussian Noise process that allows us to represent this process as a Gaussian Markov Random Field proposed by in Sørbye, Myrvoll-Nilsen and Rue (2017), and from this representation we can use Integrated Nested Laplace Approximations to perform the Bayesian estimation of param-

eters and latent volatilities. This formulation is interesting since it is computationally efficient, allowing the fast and accurate estimation of this class of models without the necessity of the use of methods based on simulation, that can become computationally prohibitive when the sample size is very great, which for example occurs in the estimation using high frequency data for realized volatility measurements.

The introduced methodology is also interesting since it allows to introduce some interesting modifications in models of long memory in the variance. In this paper we propose two extensions of the stochastic volatility model with long memory. The first extension is a two-factor model, where the first factor is the long memory component, and the second component can be thought of as a smooth variation process in the long term average of the volatility process, similar to the formulation of a Spline-Garch model (e.g., Engle and Rangel (2008)). The second modification is a stochastic volatility model with long memory adapted to the persistence patterns observed in high-frequency data. In this model we also use a structure with two latent factors, the first latent factor being a long memory process based on a Fractional Gaussian Noise representation, and a second latent factor that allows us to incorporate patterns of intraday seasonality into the stochastic volatility structure. This structure allows simultaneous estimation of the long memory process and intraday seasonality pattern, avoiding the problems associated with the two-stage estimation proposed in Deo, Hurvich and Lu (2006).

The results obtained in this work indicate that the INLA methodology allows to estimate latent parameters and volatilities in LMSV models accurately, and that the proposed extensions, such as the two-factor model and the version for high-frequency data, perform well in terms of in-sample adjustment, out-of-sample forecast, and estimation of realized volatility measurements using high-frequency data.

This paper is structured as follows. Section 6.2 exposes the formulation proposed for the LMSV model and discusses the estimation methods. Our data is described and descriptive statistics are calculated in Section 3. In Section 4 we present a Monte Carlo study of the sample properties of the estimators. In Section 5 we estimate the model for daily returns of major cryptocurrencies and compare in-sample fit and out-of-sample forecast performance to popular alternatives. An application to high frequency Bitcoin data is presented in Section 6. Section 7 concludes.

6.2 Long Memory Stochastic Volatility Models

The standard stochastic volatility (SV) model, introduced by Taylor (1986), is a log-normal mixture model in which the log normal component follows a stationary autoregressive process of first-order. Let $\{r_t\}_{t=1}^N$ be an equally spaced sequence of asset

returns, then a SV¹ model can be written as

$$r_t = \exp(h_t/2)\varepsilon_t, \quad \varepsilon_t \sim N(0, 1), \quad (6.1)$$

$$h_t = \mu + \phi(h_{t-1} - \mu) + z_t, \quad z_t \sim N(0, 1/(1 - \phi^2)\tau). \quad (6.2)$$

This model has three parameters that dictate the ARIMA(1,0,0) dynamics of latent log volatility h_t . The long run mean is μ , autoregressive persistence depends on ϕ , and the variance of latent volatility depends negatively on precision (inverse of variance) τ ².

Because log volatility is not an observed variable, maximum likelihood estimation requires evaluating a multiple integral with dimension equal to the sample size, which is computationally unfeasible. With the method of moments originally proposed by Taylor (1986) we can estimate parameters μ , ϕ , and τ , but we can not recover latent volatility h_t — which is often a quantity of interest. A traditional frequentist solution is the Quasi-Maximum Likelihood (QML) method, based on the forecast error decomposition through Kalman Filter (Harvey (1989)). Here, log squared returns are linearized and decomposed in a mean plus a first order autoregressive process, such that the model has a linear state space representation, and is then amenable to Kalman filtering. Frequentist alternatives which circumvent these issues are simulated maximum likelihood (Sandmann and Koopman (1998)), efficient method of moments (Andersen, Chung and Sørensen (1999)), empirical likelihood and generalized minimum discrepancy (Laurini and Hotta (2017)), to mention a few.

There are several Bayesian estimation methods for the class of stochastic volatility models. Because latent variables are treated as additional parameters, traditional Bayesian inference methods are directly applicable to SV models. Commonly, Markov Chain Monte Carlo algorithms are used for posterior sampling. (e.g., Gamerman and Lopes (2006) and Johanes and Polson (2005)). It is known that due to high correlation within latent variables and between latent variables and deep parameters, MCMC methods can require long chains to provide good a good posterior characterization of SV models.

Although estimation by posterior simulation methods is feasible in a wide range of practical applications, MCMC methods still can be computationally demanding, specially for models involving latent variables (whose computational burden increases with sample size). Martino et al. (2011) show how a SV model can be estimated using Integrated Nested Laplace Approximations (INLA), a class of methods introduced by Rue, Martino and Chopin (2009), which can be applied to the Bayesian estimation of parameters and latent factors of models that can be represented by, or approximated by, a Gaussian

¹ In our parameterization we directly use the marginal precision of the latent component, in contrast to the usual formulation of the precision of the innovation component of the autoregressive structure.

² Note that due to internal INLA convention we report the marginal precision of log variance h_t rather than the usual precision of iid error z_t .

Markov Random Field (GMRF). Because INLA calculations are analytic, it avoids issues inherent to posterior simulation MCMC methods, such as slow chain convergence.

Following Martino et al. (2011), if we assign mean parameter μ a Gaussian prior with zero mean and large known variance, the standard stochastic volatility model (6.1)-(6.2) can be seen as a latent Gaussian model with latent field $\mathbf{x}$:

$$\mathbf{x} = \{h_1, \dots, h_n, \mu\} \sim N(\mathbf{0}, \mathbf{Q}^{-1}(\boldsymbol{\theta}_1)), \quad (6.3)$$

where $\boldsymbol{\theta}_1 = \{\tau, \phi\}$ are parameters driving volatility. The $|\mathbf{x}| = n$ -dimensional latent Gaussian field is partially observed through the n_d conditionally independent data $\mathbf{r} = \{r_1, \dots, r_{n_d}\}$ with likelihood

$$\pi(\mathbf{r}|\mathbf{x}, \boldsymbol{\theta}_2) = \prod_{t=1}^{n_d} \pi(r_t|\mathbf{h}, \boldsymbol{\theta}_2), \quad (6.4)$$

where $\boldsymbol{\theta}_2$ are parameters of the ε_t return process, which in our case is a precision of one, since $\varepsilon_t \sim N(0, 1)$.

Let $\boldsymbol{\theta} = \{\boldsymbol{\theta}_1, \boldsymbol{\theta}_2\}$. The main goal in estimating SV models is evaluating the marginal distributions

$$\pi(\mathbf{x}, \boldsymbol{\theta}|\mathbf{r}) \propto \pi(\boldsymbol{\theta})\pi(\mathbf{x}|\boldsymbol{\theta}) \prod_{t=1}^n \pi(r_t|h_t, \boldsymbol{\theta}). \quad (6.5)$$

The INLA procedure is a computationally efficient method to compute these marginal distributions. At its core is a very fast Gaussian approximation to densities of the form

$$\pi(\mathbf{x}|\mathbf{r}, \boldsymbol{\theta}) \propto \exp \left\{ -\frac{1}{2} \mathbf{x}^T \mathbf{Q} \mathbf{x} + \sum g_t(h_t) \right\}, \quad (6.6)$$

where $g_t(h_t) = \log \pi(r_t|h_t, \boldsymbol{\theta})$. The Gaussian approximation $\bar{\pi}_G(\mathbf{x}|\mathbf{r}, \boldsymbol{\theta})$ is obtained by matching the curvature at the posterior mode value $\mathbf{m}$, which is computed iteratively using a Newton-Raphson algorithm. That is,

$$\bar{\pi}_G(\mathbf{x}|\mathbf{r}, \boldsymbol{\theta}) = K_1 \exp \left\{ -\frac{1}{2} (\mathbf{x} - \mathbf{m})^T (\mathbf{Q} + \text{diag}(\mathbf{c})) (\mathbf{x} - \mathbf{m}) \right\}, \quad (6.7)$$

where K_1 is a normalizing constant, and the vector $\mathbf{c}$ is a correction term given by second-order terms in the Taylor expansion of $\sum g_t(h_t)$ at the modal value $\mathbf{m}$.

6.2.1 Long Memory Stochastic Volatility Models

Although SV models capture key aspects of returns series such as almost no autocorrelation in levels and excess kurtosis, for some assets the decay in autocorrelation of squared returns can be slower than implied by the autoregressive structure in (6.2).

An alternative is to model log volatility as a stationary process with long memory. A weakly stationary process has long memory if its autocovariance function $\gamma(k)$ for long k lags satisfy

$$\gamma(k) \sim C_1 k^{H-1}, \quad \text{as } k \rightarrow \infty, \quad (6.8)$$

for $C_1 > 0$. Or, in the frequency domain, if its spectral density $f(\omega)$ for frequencies $\omega \in [-\pi, \pi]$ close to zero satisfy

$$f(\omega) \sim C_2 |\omega|^{-2H+1}, \quad \text{as } \omega \rightarrow 0, \quad (6.9)$$

for $C_2 > 0$ (Beran (2017)).

If the log volatility component h_t has long range dependence properties then we have the long memory stochastic volatility model (LMSV) as introduced by Harvey (1998) and Breidt, Crato and Lima (1998).

The LMSV model structure used in this article is given by:

$$y_t = \exp(h_t/2)\varepsilon_t, \quad \varepsilon_t \sim N(0, 1), \quad (6.10)$$

$$h_t = \mu + fGn(H, \tau), \quad (6.11)$$

where μ is a mean parameter, h_t is a Fractional Gaussian Noise (fGn) process (Mandelbrot and Ness (1968)), defined by its autocorrelation function

$$\gamma_y(k) = \frac{1}{2} \left(|k-1|^{2H} - 2|k|^{2H} + |k+1|^{2H} \right), \quad (6.12)$$

for $k = 0, 1, \dots, n-1$, with marginal precision τ . Notice the autocorrelation function is hyperbolic decaying, as $\gamma_y(k) \sim H(2H-1)k^{2(H-1)}$ when $k \rightarrow \infty$.

The LMSV model has three parameters, which characterize log volatility dynamics. The long run mean of log-variance is μ , the marginal precision of h_t is given by τ , and temporal persistence of the fGn component is dictated by the Hurst exponent parameter H .

A popular alternative approach with LMSVs is to model log variance as an ARFIMA process. ARFIMA and fGn are closely related, specially if the autoregressive and moving average order of the ARFIMA are 0. Conceptually, a fGn comes from the fractional differentiation of a continuous Brownian motion (Hosking (1981)), while ARFIMA comes from the fractional difference of a discrete ARIMA process.

Because long memory processes do not have a finite state representation, estimation of LMSV models is not straightforward. Note many of the estimation methods for SV models we mention take advantage of the Markovian structure of some sort of state space representation of the model. Estimation of LMSV models has traditionally been

done in the time domain by simulation methods and in the frequency domain by spectral likelihood estimators. Simulation-based Bayesian estimation and Quasi-Maximum likelihood estimations of LMSV are based on truncated state space representations of the long memory process, as discussed in Chan and Palma (1998), Chan and Petris (2000) and Ferraz and Hotta (2007). A popular approach is to deal with the approximate representation of an ARFIMA process as a long lagged ARMA model. A simple representation is to perform the expansion of an ARFIMA(0, d , 0) through an autoregressive process of the form:

$$\sum_{i=0}^P \Pi_i h_{t-i} = \eta_t \quad (6.13)$$

with $\Pi_0 = 1$, $\Pi_j = [(j-d-1)/j]\Pi_{j-1}$, P the order of approximation, h_t the latent process and η_t a white noise process. This representation is quite interesting since it is enough to rewrite the representation in state space to that of an AR(P) model, in this case with coefficients determined by the above expansion. The usual Quasi-Maximum likelihood and MCMC estimators can be adapted to this representation.

6.2.2 Gaussian Markov Random Field approximation of Fractional Gaussian Noise

Sørbye, Myrvoll-Nilsen and Rue (2017) take advantage of the relationship between long memory and aggregation (Granger (1980), Beran, Schützner and Ghosh (2010)) to construct an Gaussian Markov Random Field based approximation to a fractional Gaussian noise model from weighted sums of independent AR(1) components. They propose to match the autocorrelation function of this composite autoregressive process to the autocorrelation (6.12) of a fGn.

Following Sørbye, Myrvoll-Nilsen and Rue (2017), consider m independent AR(1) processes

$$z_{j,t} = \phi_j z_{j,t-1} + \xi_{j,t}, \quad j = 1, \dots, m, \quad t = 1, \dots, n, \quad (6.14)$$

where $0 < \phi_j < 1$ is the first-order autoregressive parameter of the j -th process. Also, let $\{\xi_{j,t}\}_{j=1}^m$ be zero-mean independent Gaussian shocks with variance $\sigma_{\xi,j}^2 = 1 - \phi_j^2$. Define the cross-sectional aggregation of the m processes as

$$\bar{z}_m = \sigma \sum_{j=1}^m \sqrt{w_j} z^{(j)}, \quad (6.15)$$

where $z^{(j)}$ denotes $(z_{j,1}, z_{j,2}, \dots, z_{j,n})$ and the weights w_j sum to one. Haldrup and Valdés (2017) study the finite sample properties of similar aggregations of AR(1) processes.

The autocorrelation function of (6.15) is simply

$$\gamma_{\bar{z}_m}(k) = \sum_{j=1}^m w_j \phi_j^{|k|}, \quad k = 0, 1, \dots, n-1. \quad (6.16)$$

The idea put forward by Sørbye, Myrvoll-Nilsen and Rue (2017) is to fit weights $\mathbf{w} = \{w_j\}_{j=1}^m$ and coefficients $\boldsymbol{\phi} = \{\phi_j\}_{j=1}^m$ such that the autocorrelation function (6.16) of the composite AR(1) process matches the autocorrelation function of a true fGn process as in (6.12). Values of $(\mathbf{w}, \boldsymbol{\phi})$ are obtained by minimizing the squared error

$$(\mathbf{w}, \boldsymbol{\phi})_H = \underset{(\mathbf{w}, \boldsymbol{\phi})}{\operatorname{argmin}} \sum_{k=1}^{k_{\max}} \frac{1}{k} (\gamma_{\bar{z}_m}(k) - \gamma_y(k))^2, \quad (6.17)$$

where $k_{\max}$ is an arbitrary upper limit to the number of lags included. Since the squared error is weighted by k , persistence at very distant lags have little impact to the objective function (6.17).

With this approximation of the fGn process we can represent the LMSV model (6.10)-(6.11) as a latent GMRF, and estimate it using the INLA method, as shown by Martino et al. (2011). Following the recommendations in Sørbye, Myrvoll-Nilsen and Rue (2017), we use a third order approximation to represent the fGn process. Further details on the implementation can be found in Sørbye, Myrvoll-Nilsen and Rue (2017).

6.2.3 Alternative Specifications

The INLA methodology is not restrictive and SV models can be easily augmented with additional factors. As long as the affine structure of log volatility is preserved, the model has a GMRF representation similar to equation (6.3), and can be estimated in the same manner. We consider alternative formulations for both SV and LMSV models, in which log volatility is composed by a smooth spline trend in addition to either the AR(1) or fGn processes. These specifications are similar to the Spline-GARCH model of Engle and Rangel (2008) in which conditional volatility is subject to low frequency variations, which work as a time varying long run average.

The AR(1) Spline-SV model can be written as

$$r_t = \exp(h_t + s_t)/2 \varepsilon_t, \quad \varepsilon_t \sim N(0, 1), \quad (6.18)$$

$$h_t = \phi h_{t-1} + z_t, \quad z_t \sim N(0, 1/(1 - \phi^2)\tau_h), \quad (6.19)$$

where s_t follows a second-order random walk, which is constructed by assuming independent second-order increments with precision τ_s . That is,

$$\Delta^2 s_t = s_t - 2s_{t+1} + s_{t+2}, \quad s_t \sim N(0, 1/\tau_s). \quad (6.20)$$

The use of a second-order random walk structure to represent a spline process is common in statistics, see, for example, Lindgren and Rue (2008). Note that an analogous relationship exists between the popular Hodrick-Prescott (HP) filter and spline representations, as discussed in Harvey and Jaeger (1993), Harvey and Trimbur (2008) and Paige and Trindade (2010).

Since we use an unrestricted specification for this component, i.e, the component does not sum to zero, it can be used to capture a long-run average of the log-variance that is varying smoothly. Hence in this specification it is not necessary to include the μ mean parameter, already captured by the dynamics of this process.

Therefore, the AR(1) Spline model has three parameters: the autoregressive coefficient ϕ , the marginal precision of log volatility τ_h , and the precision of spline (trend) component τ_s . More details on the specification of the second-order random walk model can be found in the Appendix.

Likewise, consider the LMSV Spline model

$$r_t = \exp((h_t + s_t)/2)\varepsilon_t, \quad \varepsilon_t \sim N(0, 1), \quad (6.21)$$

$$h_t = fGn(H, \tau), \quad (6.22)$$

$$\Delta^2 s_t = s_t - 2s_{t+1} + s_{t+2}, \quad s_t \sim N(0, 1/\tau_s). \quad (6.23)$$

This model has three parameters. The Hurst exponent H , which determines the temporal persistence of the fGn process, the marginal precision of fGn component τ_h , and the precision τ_s of the second-order random walk spline.

Another alternative form introduced in this work is the addition of a seasonal (periodic) component $seas_t$ in the conditional variance structure. In this specification we introduce an additional stochastic process with periodicity M , with the restriction that the sum of the m components is zero:

$$r_t = \exp(h_t + seas_t)/2\varepsilon_t, \quad \varepsilon_t \sim N(0, 1), \quad (6.24)$$

$$h_t = \mu + fGn(H, \tau), \quad (6.25)$$

$$seas_t = seas_{t+1} + \dots + seas_{t+m+1} + \nu_t, \quad \nu_t \sim N(0, 1/\tau_{seas}). \quad (6.26)$$

This component is important in the modeling of intraday returns, where there is the presence of periodic patterns in the volatility structure, as discussed in Deo, Hurvich and Lu (2006). The period m depends on the aggregation used in the estimation with high frequency returns. For example in section 7 we use returns of 10 minutes, and this defines the number of components used in the representation of intraday periodicity.

The direct incorporation of this periodicity structure is interesting since it avoids the use of multi-stage methods for the estimation of the LMSV model for high frequency

data. For example in Deo, Hurvich and Lu (2006) the periodicity pattern is estimated in a first stage, and then returns adjusted for this pattern are modeled as an LMSV process. Our formulation allows the simultaneous estimation of these components, avoiding problems associated with estimation in stages.

6.3 Monte Carlo Simulation Study

We now perform Monte Carlo experiments to evaluate the small sample properties of the INLA fGn approximation and estimation of LMSV models described in Section 6.2. We compare INLA estimates with estimates from a MCMC simulation algorithm based on a truncated state space representation of long memory processes proposed by Chan and Petris (2000). The long memory process is approximated as an AR(10) process. For a detailed description of this method see Chan and Petris (2000) or Ferraz and Hotta (2007).

Two parameter vectors for the data generating process are considered. The first one, $\{\mu = -9.90, H = 0.931, \tau = 0.962\}$ correspond to the parameters estimated from daily S&P500 returns, and represents traditional market conditions. The second parameter vector, $\{\mu = -7.182, H = 0.938, \tau = 0.294\}$, comes from point estimates from a series of Bitcoin daily returns, and thus represents an unconventional market with substantially higher levels of observed volatility.

For each parameter vector we simulate 1000 samples of LMSV model (6.10)-(6.2) with 500 and 1000 observations. In each experiment we compare estimated to true parameter values in terms of mean error (me), root mean squared error (rmse), and mean absolute error (mae). In order to avoid issues with initial values, each simulated sample includes an initial period of length 300, which is discarded. In general, the results between the different parameter vectors and sample sizes are consistent. It can be observed that the estimation by INLA presents better results in terms of me, rmse and mae for the parameters μ and τ , except for the mean error of the parameter μ for the parameter vector with greater unconditional volatility (Tables 6.3 and 6.4). The MCMC algorithm performs better in estimating the H parameter in all experiments. However, in observing the adjustment to the true latent volatility process, the INLA-based methodology presents a much superior performance in terms of me, rmse and mae in all parameter configurations and sample sizes analyzed. This result is very relevant, since the main objective for risk measurement is to accurately recover the true volatility of the process, which is the necessary input for the calculation of risk management measures such as Value at Risk. In this aspect the proposed methodology has a satisfactory performance for empirical applications, and showing the validity of the method for real problems of risk measurement and management.

	me	rmse	mae
μ MCMC	-0.0157	0.216	0.175
μ INLA	-0.000175	0.202	0.163
H MCMC	-0.0199	0.0941	0.0537
H INLA	-0.0861	0.111	0.0876
τ MCMC	0.747	0.785	0.747
τ INLA	0.764	1.07	0.792
h_t MCMC	0.00364	0.00364	0.00277
h_t INLA	-0.000447	0.00244	0.00179

Table 6.1 – Monte Carlo Experiment LMSV Model - First parameter set ($\mu=-9.9$, $H=0.931$, $\tau=0.962$) - Sample size 500

	me	rmse	mae
μ MCMC	-0.0105	0.119	0.0959
μ INLA	-6.52e-05	0.113	0.0908
H MCMC	0.0271	0.0456	0.0399
H INLA	-0.0435	0.0639	0.0476
τ MCMC	0.742	0.763	0.742
τ INLA	0.413	0.589	0.458
h_t MCMC	0.00365	0.00365	0.00278
h_t INLA	-0.000424	0.00243	0.00178

Table 6.2 – Monte Carlo Experiment LMSV Model - First parameter set ($\mu=-9.9$, $H=0.931$, $\tau=0.962$) - Sample size 1000

	me	rmse	mae
μ MCMC	-0.0153	0.368	0.299
μ INLA	0.0705	0.352	0.284
H MCMC	0.0164	0.0394	0.033
H INLA	-0.0913	0.0959	0.0913
τ MCMC	0.715	0.722	0.715
τ INLA	0.6	0.639	0.601
h_t MCMC	0.0325	0.0325	0.0225
h_t INLA	-0.00411	0.0188	0.0118

Table 6.3 – Monte Carlo Experiment LMSV Model - Second parameter set ($\mu=-7.182$, $H=0.938$, $\tau=0.294$) - Sample size 500

6.4 Data

To check the fit properties of the proposed models for real data, we choose a set with four assets. We consider the three largest cryptocurrencies as of march 2019: Bitcoin (btc), Ethereum (eth), Litecoin (ltc). We use daily data from 2015-01-01 until 2019-03-10, corresponding to 1529 observations. In the Ethereum estimations, we discard the first 7 observations due to extreme values observed in this first week. As discussed in the introduction of the article, there is already ample evidence of the presence of long

	me	rmse	mae
μ MCMC	-0.011	0.196	0.157
μ INLA	0.0352	0.186	0.148
H MCMC	0.0369	0.0425	0.0401
H INLA	-0.0689	0.072	0.0689
τ MCMC	0.707	0.711	0.707
τ INLA	0.363	0.378	0.363
h_t MCMC	0.0328	0.0328	0.0227
h_t INLA	-0.00438	0.019	0.0117

Table 6.4 – Monte Carlo Experiment LMSV Model - Second parameter set ($\mu=-7.182$, $H=0.938$, $\tau=0.294$) - Sample size 1000

memory in these cryptocurrencies.

For the sake of context we also repeat our analysis to the S&P500 (sp500) index. The sample used contains daily returns from 2012-01-03 until 2018-12-31, corresponding to 1759 observations.

	btc	eth	ltc	sp500
Minimum	-0.320	-0.519	-0.236	-0.042
Quartile 1	-0.027	-0.018	-0.011	-0.003
Median	-0.001	0.000	0.002	0.000
Arithmetic Mean	0.004	0.002	0.002	0.000
Geometric Mean	0.002	0.000	0.001	0.000
Quartile 3	0.031	0.019	0.017	0.005
Maximum	0.404	0.518	0.224	0.048
Variance	0.004	0.004	0.002	0.000
Stdev	0.067	0.060	0.039	0.008
Skewness	0.446	0.664	-0.336	-0.423
Kurtosis	4.290	13.727	5.419	3.282

Table 6.5 – Descriptive Statistics - Bitcoin (btc), Ethereum (eth), Litecoin (ltc) and S&P 500 (sp500)

Table 6.5 reports descriptive statistics of daily log returns of the three cryptocurrencies Bitcoin, Ethereum, Litecoin, and the SP&500 index. Figure 6.1 plots the trajectory of returns. We can observe that the cryptocurrencies have a high volatility and have asymmetry coefficients distant from zero and high kurtosis, in contrast to the more typical behavior presented by the S&P 500. This behavior is consistent with the great extremes observed in this period, and also may be related to higher persistence patterns in volatility.

Figure 6.1 – Returns - Bitcoin (btc), Ethereum (eth), Litecoin (ltc), S&P 500 (sp500)

6.5 Estimation Results - Daily Data

We now report and discuss estimation results using INLA for daily returns for the four assets in our sample. Prior specification and implementation details are reported in the Appendix. For each asset we use four specifications. The first is a traditional SV model (Ar1SV) using an AR(1) structure for the dynamics of the conditional variance. The second specification is based on an LMSV model with variance persistence given by a fractional Gaussian noise process. The next two models add a second latent factor in the standard SV and LMSV models, using a spline component to model the mean of the volatility in time using a structure of second order random walk as described in Section 2.

The posterior distribution for the parameters estimated for the four assets is shown in the Tables 6.6-6.9. We can observe some common results among the various specifications. The parameters H and ϕ in models Ar1SV and LMSV are high, consistent with the persistence patterns observed in financial time series. In the specifications with the spline component, the results seem to indicate that the H coefficient is relevant, but in the Ar1SV Spline specifications the coefficient ϕ has a high dispersion, indicating that only the spline component is required in this specification. The LMSV plus spline decomposition seems to be an interesting specification for most of the series analyzed. In the pure Ar1SV and LMSV models the values of μ are consistent with the estimated unconditional volatility in all assets, indicating higher mean volatilities for the cryptocurrencies, and a

smaller mean volatility for sp500. The precision parameters τ are also consistent with the observed variation in the trajectory of asset returns.

	mean	sd	0.025quant	0.5quant	0.975quant
Ar1SV					
μ	-7.32	0.195	-7.7	-7.32	-6.93
ϕ	0.923	0.0172	0.885	0.924	0.952
τ_h	0.506	0.0724	0.379	0.501	0.663
LMSV					
μ	-7.44	1.3	-10.1	-7.45	-4.77
H	0.93	0.0224	0.879	0.932	0.966
τ_h	0.326	0.0664	0.214	0.32	0.473
Ar1SV Spline					
μ					
ϕ	-0.491	0.494	-0.992	-0.668	0.717
τ_h	1.8e+04	1.8e+04	1.12e+03	1.26e+04	6.58e+04
τ_s spline	1.33e+03	445	674	1.26e+03	2.4e+03
LMSV Spline					
μ					
H	0.849	0.0436	0.755	0.853	0.924
τ_h	0.571	0.106	0.394	0.561	0.809
τ_s spline	7.67e+04	3.51e+04	2.82e+04	7.02e+04	1.63e+05

Table 6.6 – Posterior distribution estimated parameters - Bitcoin (btc)

In Table 6.10 we show the estimated marginal likelihood for each model, as well as a measure of model comparison, the Waic. Waic, proposed in Vehtari, Gelman and Gabry (2016), is an adequate information criterion for the comparison of Bayesian models of different complexities, and can be interpreted in a way analogous to the traditional information criteria. In this table the values in bold show the best specification chosen by the Waic criterion for each asset. For the three cryptocurrencies, the LMSV spline model is chosen as the best specification, while for the S&P 500 the LMSV model is selected. These results seem to indicate that the LMSV specification is suitable for all series, and for the cryptocurrencies the inclusion of a smooth change component in the unconditional average seems to better capture the evolution of this market. The inclusion of this component is a way of incorporating a change of parameters in the unconditional volatility process, and these results are consistent with the many changes that occurred in these markets in the analyzed period. To show the dynamics of this component, we show in Figure 6.2 the estimated Spline component for the Bitcoin returns. We can observe that this component adjusts the level variations of the latent volatility, and seems to accompany the possible permanent changes in the latent volatility of this asset.

For the in-sample adjustment of the different specifications tested, we show in Table 6.11 some traditional measures of adjustment (mean error, root mean squared error

	mean	sd	0.025quant	0.5quant	0.975quant
Ar1SV					
μ	-6.08	0.144	-6.36	-6.08	-5.79
ϕ	0.868	0.0277	0.807	0.87	0.916
τ_h	0.658	0.0919	0.497	0.651	0.858
LMSV					
μ	-6.09	1.29	-8.74	-6.09	-3.41
H	0.927	0.0269	0.865	0.931	0.969
τ_h	0.367	0.0881	0.223	0.357	0.567
Ar1SV Spline					
μ					
ϕ	-0.493	0.497	-0.992	-0.675	0.722
τ_h	1.82e+04	1.82e+04	1.17e+03	1.28e+04	6.65e+04
τ_s spline	1.16e+03	453	531	1.07e+03	2.28e+03
LMSV Spline					
μ					
H	0.922	0.0307	0.851	0.926	0.969
τ_h	0.397	0.115	0.214	0.383	0.664
τ_s spline	9.42e+04	4.3e+04	3.52e+04	8.62e+04	2e+05

Table 6.7 – Posterior distribution estimated parameters - Ethereum (eth)

(rmse) and mean absolute error (mae)). We compared the conditional volatility fitted for each model with absolute returns as proxy for true unobserved volatility.

We can observe that the LMSV Spline model is chosen by all criteria as the best model in terms of in-sample fit for Bitcoin, whereas for Ethereum the LMSV model is chosen by rmse and mae. For Litecoin the model LMSV Spline is chosen by the me and mae criteria, and LMSV by the rmse. For S&P 500 the LMSV is chosen by the rmse and mae criteria, and the LMSV by me. In terms of in-sample fit, the results obtained seem to indicate the presence of a long memory component in the volatility process. An out-of-sample forecast analysis is presented in the next section.

In the Figures 6.3 - 6.6 we show a comparison of the volatility estimated by each specification, compared to the absolute returns. The results seem to indicate that the adjustment of the estimated models is consistent with the pattern of volatility observed in all the analyzed series.

In the same Figures 6.3 - 6.6 we also show dynamic Value at Risk (VaR) measurements calculated using the different volatility components estimated by the four analyzed specifications. VaR is estimated using the approximation $\text{VaR}(\alpha) = \hat{y} - q_\alpha * \sigma_t$, where $\hat{y}$ is the sample mean of the process, q_α is the quantile α of a standard Normal distribution and σ_t is the estimated conditional volatility. In this experiment we show the results for the calculation of a VaR with $\alpha = 0.05$, a usual measure of tail risk. Results for other VaR levels are available from the authors.

	mean	sd	0.025quant	0.5quant	0.975quant
Ar1SV					
μ	-6.85	0.185	-7.21	-6.85	-6.48
ϕ	0.878	0.0194	0.837	0.879	0.913
τ_h	0.329	0.0385	0.259	0.327	0.41
LMSV					
μ	-6.77	1.54	-9.9	-6.78	-3.59
H	0.931	0.0216	0.883	0.933	0.966
τ_h	0.241	0.0533	0.151	0.236	0.358
Ar1SV Spline					
μ					
ϕ	-0.492	0.495	-0.992	-0.671	0.718
τ_h	1.75e+04	1.79e+04	1.02e+03	1.21e+04	6.52e+04
τ_s spline	322	139	139	293	672
LMSV Spline					
μ					
H	0.878	0.0361	0.801	0.88	0.941
τ_h	0.376	0.0772	0.242	0.37	0.544
τ_s spline	7.85e+04	3.63e+04	3.03e+04	7.12e+04	1.69e+05

Table 6.8 – Posterior distribution estimated parameters - Litecoin (ltc)

In general, dynamic VaR measurements appear to follow the observed tail risk for all assets. In order to verify more carefully the performance of the different specifications for conditional volatility, we show in the Table 6.12 the proportion of observed violations within the sample (prop. viol.), and the p-values from the backtesting method for the VaR proposed by Christoffersen (1998) to verify if the empirical coverage of the estimated VaR is statistically equal to the VaR nominal value, against an alternative hypothesis of empirical coverage distinct from the nominal, indicating a problem in the VaR estimation.

In this analysis the results obtained are quite different in relation to the adjustment measures analyzed above. The proportion of violations obtained by the specifications using the LMSV model indicate a number of violations less than the expected nominal value, indicating that the LMSV based models overestimate the tail risk. The specification based on the Ar(1) structure seems to indicate an adequate level of coverage, indicating that for the tail risk measurement in the analyzed period the traditional Ar(1) model seems to perform better.

An important issue is that in this work we compare the various specifications through comparisons of information criteria, in-sample and out-of-sample adjustment measures, as well as derived measures such as Value at Risk. However, it is possible to perform direct comparison procedures between Ar1SV and LMSV models using marginal likelihoods and Bayes factor through a structure of priors that allows an equivalence between the Ar(1) and H coefficients in this two specifications. A simple way of performing

	mean	sd	0.025quant	0.5quant	0.975quant
Ar1SV					
μ	-10.1	0.127	-10.3	-10.1	-9.81
ϕ	0.928	0.0162	0.892	0.929	0.956
τ_h	1.15	0.174	0.846	1.14	1.53
LMSV					
μ	-10	1.51	-13.1	-10	-6.88
H	0.959	0.0217	0.905	0.963	0.988
τ_h	0.467	0.178	0.205	0.438	0.894
Ar1SV Spline					
μ					
ϕ	-0.488	0.498	-0.992	-0.667	0.724
τ_h	1.81e+04	1.81e+04	1.15e+03	1.27e+04	6.63e+04
τ_s spline	2.28e+04	1.1e+04	7.8e+03	2.07e+04	5.01e+04
LMSV Spline					
μ					
H	0.922	0.0342	0.842	0.927	0.973
τ_h	0.866	0.289	0.418	0.828	1.54
τ_s spline	1.18e+05	4.6e+04	5.06e+04	1.11e+05	2.28e+05

Table 6.9 – Posterior distribution estimated parameters - S&P 500 (sp500)

	btc		eth		ltc		sp500	
	mlik	waic	mlik	waic	mlik	waic	mlik	waic
Ar1SV	3179.3	-6533.9	1889.3	-3948.8	2705.2	-5730.2	6188.1	-12517
LMSV	3214.6	-6659.4	1903.3	-3994.2	2743.1	-5810.5	6196.1	-12531
Ar1SV Spline	3093.5	-6350.9	1804.9	-3771.7	2540.7	5289.2	-6120.9	-12348
LMSV Spline	3200.2	-6671.6	1884.0	-3994.5	2727.8	-5813.4	6175.1	-12528

Table 6.10 – Model Fit Measures - Marginal Likelihood and Waic

	btc			eth			ltc			sp500		
	me	rmse	mae	me	rmse	mae	me	rmse	mae	me	rmse	mae
Ar1SV	0.00483	0.0226	0.0162	0.00701	0.0369	0.0275	0.00421	0.0327	0.0205	0.00117	0.00478	0.00371
LMSV	0.00341	0.0183	0.0129	0.00596	0.0317	0.0235	0.00375	0.0275	0.0171	0.00107	0.00434	0.00338
Ar1SV Spline	0.00792	0.0268	0.02	0.0137	0.0465	0.0359	0.0122	0.044	0.0299	0.00174	0.00556	0.00439
LMSV spline	0.00282	0.0182	0.0128	0.00566	0.032	0.0237	0.00285	0.0276	0.0170	0.00106	0.00442	0.00342

Table 6.11 – In Sample Error Measures

this procedure is through the structure of penalized complexity (PC) priors, introduced in Simpson et al. (2017) and adapted for the case of fractional Gaussian Noise processes in Sørbye and Rue (2018). In this work we use the PC-priors structure for the coefficients of the Ar1 and LMSV models, and thus we could compare models using these structures, but we prefer to focus on comparisons in terms of fit results and applications.

Figure 6.2 – Spline Component - Bitcoin

	btc		eth		ltc		sp500	
	prop. viol.	p.value	prop. viol.	p.value	prop. viol.	p.value	prop. viol.	p.value
Ar1SV	0.0549	0.378	0.0484	0.849	0.0419	0.159	0.0455	0.412
LMSV	0.0314	0.0005	0.0284	0.0001	0.0222	4.02e-08	0.033	0.0006
Ar1SV Spline	0.0576	0.177	0.0553	0.373	0.0477	0.725	0.0517	0.743
LMSV Spline	0.0334	0.0022	0.0354	0.0131	0.0242	5.23e-07	0.0358	0.0052

Table 6.12 – VaR statistics

6.6 Out-of-sample forecasting

The previous analyzes showed the in-sample performance of the conditional volatility models. To verify the predictive performance for conditional volatility out-of-sample, we performed a forecast experiment using a rolling sample for the last 22 observations in the sample, which are left out of the estimation procedure.

In this analysis we used rolling sample with size equal to the size of the series minus 22, and with this series we estimate the four specifications analyzed above. From these estimates, we predict future volatility for 1, 5 and 22 steps-ahead, and compare it with the observed values of absolute returns (proxy for true unobserved volatility). Then we added one more observation in the estimation sample, and replicated the forecasting process until the end of this augmented sample. The prediction results for each series analyzed, four different specifications and the three forecast horizons are presented in the Tables 6.13-6.14. In this table we also add as predictive performance measures the mean percent error (MPE) and the mean absolute percent error (MAPE). The best results in

Figure 6.3 – Bitcoin (btc) - Fitted Volatility and VaR (5%)

each category are marked in bold.

The out-of-sample forecast results shown in the Tables 6.13-6.14 indicate certain variability in the models with the best performance in each forecast horizon, but generally indicate a better predictive performance for the two LMSV-based models. In terms of average error, LMSV-based models obtain the best performance across all series and forecast horizons. Looking at the RMSE and MAE, there is a balance of performance, with very close performances between all the different specifications analyzed. In terms of MPE and MAPE the results indicate that the LMSV models have the best performance for most of the series and forecast horizons analyzed.

Although these results can not be generalized to other assets or sample periods, the results indicate that models with long-memory structures in the conditional volatility

Figure 6.4 – Ethereum (etc) - Fitted Volatility and VaR (5%)

perform well in terms of out-of-sample forecasting, and support the practical use of the introduced methods in this work.

6.7 Application to high frequency data

A natural application of long memory models is to model conditional volatility in high frequency return data. This type of application is important since it is possible to use the estimated conditional variance for high frequency data to calculate realized volatility measurements (e.g. Maasoumi and McAleer (2008), McAleer and Medeiros (2008) and Lieberman and Phillips (2008)) for daily data. As discussed in Deo, Hurvich and Lu (2006), models of stochastic volatility with long memory are particularly important in

Figure 6.5 – Litecoin (ltc) - Fitted Volatility and VaR (5%)

high frequency data, since in this context the presence of long memory is much more frequent.

This method is particularly interesting since it directly incorporates the observed dependence in the latent volatility process and is an alternative form to the calculation of realized volatility measures based on robust estimators for the presence of dependence and market microstructure noise.

The estimation methodology based on INLA and the approximation of a fGn by a Gaussian Markov Random Field is especially interesting. It is computationally efficient in terms of computational speed and memory usage, and thus can be used for the high number of observations observed in high frequency data. The second advantage is the possibility of directly incorporating into the model a latent variable to capture patterns

Figure 6.6 – S&P 500 (sp500) - Fitted Volatility and VaR (5%)

of intraday seasonality in market volatility, without the need for multi-step estimation as proposed in Deo, Hurvich and Lu (2006).

We estimate a LM SV model using 10-minute intraday Bitcoin data from Bitstamp exchange for the period 2017-06-27 until 2018-06-27, corresponding to a sample with 52417 intraday observations. Figure 6.7 shows the intraday Bitcoin returns analyzed. Based on this estimation, we compute the realized volatility for each day by aggregating the estimated conditional volatilities for each day, and compared with a robust volatility measure to market microstructure noise calculated using the Two Time Scale realized variance methodology proposed by Zhang and Mykland P. A. Ait-Sahalia (2005).

Table 6.15 shows the estimated parameters using the LM SV model for this series, and Figure 6.9 shows the conditional volatility estimated by the model for the intraday

	Bitcoin	ME	RMSE	MAE	MPE	MAPE
Ar1SV	1-step	0.007	0.017	0.013	36.903	75.409
LMSV	1-step	0.007	0.017	0.014	40.740	75.776
Ar1SV Spline	1-step	0.001	0.016	0.009	5.658	80.505
LMSV Spine	1-step	0.004	0.016	0.011	25.249	76.647
Ar1SV	5-step	0.006	0.021	0.016	22.469	86.748
LMSV	5-step	0.007	0.020	0.016	33.256	81.065
Ar1SV Spline	5-step	-0.003	0.019	0.010	-41.152	108.210
LMSV Spine	5-step	0.001	0.019	0.012	-1.068	89.144
Ar1SV	22-step	0.008	0.022	0.019	33.418	79.558
LMSV	22-step	0.007	0.022	0.018	29.386	80.877
Ar1SV Spline	22-step	-0.010	0.023	0.012	-175.393	224.260
LMSV Spine	22-step	-0.004	0.022	0.013	-56.994	133.389
	Ethereum	ME	RMSE	MAE	MPE	MAPE
Ar1SV	1-step	0.017	0.026	0.022	44.443	61.793
LMSV	1-step	0.019	0.026	0.022	49.578	60.612
Ar1SV Spline	1-step	0.019	0.028	0.023	44.164	61.706
LMSV Spine	1-step	0.009	0.024	0.018	6.272	75.008
Ar1SV	5-step	0.017	0.032	0.027	45.238	66.175
LMSV	5-step	0.017	0.032	0.027	45.326	66.290
Ar1SV Spline	5-step	0.019	0.033	0.027	46.341	63.711
LMSV Spine	5-step	0.000	0.028	0.019	-15.373	95.000
Ar1SV	22-step	0.013	0.045	0.039	28.284	81.973
LMSV	22-step	0.009	0.044	0.037	20.115	86.068
Ar1SV Spline	22-step	0.011	0.048	0.039	15.111	90.966
LMSV Spine	22-step	-0.021	0.051	0.032	-457.784	509.686

Table 6.13 – Out of Sample Forecast Measures: Bitcoin and Ethereum

data compared to the absolute intraday returns. We can observe that the model can adjust the temporal variation observed in the intraday returns.

The patterns of intraday seasonality estimated by the model can be seen in Figure 6.8, which shows that the seasonality has a substantial variation during the day.

From the estimated intraday conditional variance, we construct a measure of realized volatility by aggregating these results for each day, summing the estimated conditional variances. In figure 6.10 we show the daily volatility measure constructed from the estimated LMSV model, compared to the measure of realized volatility calculated using intraday returns by the Two-Time Scale realized variance method. We can observe that the volatility constructed through the LMSV model is very close to the measure of realized variance robust to market microstructure noise.

Figure 6.7 – 10-minutes intraday Bitcoin returns

Figure 6.8 – Fitted Intraday Volatility Seasonality

Figure 6.9 – Absolute 10-minute Returns and Fitted Intraday Volatility

Figure 6.10 – Realized Variance and LMSV-based Realized Variance

	Litecoin	ME	RMSE	MAE	MPE	MAPE
Ar1SV	1-step	0.001	0.059	0.035	-36.606	109.256
LMSV	1-step	0.001	0.056	0.032	-19.319	94.812
Ar1SV Spline	1-step	-0.001	0.057	0.032	-34.121	107.418
LMSV Spine	1-step	0.000	0.056	0.032	-28.522	103.568
Ar1SV	5-step	-0.006	0.056	0.033	-36.739	106.149
LMSV	5-step	-0.004	0.054	0.032	-19.319	92.489
Ar1SV Spline	5-step	-0.009	0.056	0.032	-50.927	116.654
LMSV Spine	5-step	-0.007	0.056	0.033	-37.072	107.497
Ar1SV	22-step	-0.008	0.038	0.025	-24.611	76.615
LMSV	22-step	-0.003	0.038	0.026	-8.238	68.537
Ar1SV Spline	22-step	-0.014	0.041	0.028	-70.668	112.848
LMSV Spine	22-step	-0.008	0.040	0.027	-37.054	86.990

	S&P500	ME	RMSE	MAE	MPE	MAPE
Ar1SV	1-step	0.001	0.059	0.035	-36.605	109.256
LMSV	1-step	0.001	0.056	0.032	-19.320	94.813
Ar1SV Spline	1-step	-0.001	0.056	0.032	-34.084	107.365
LMSV Spine	1-step	0.000	0.056	0.032	-28.525	103.568
Ar1SV	5-step	-0.006	0.056	0.033	-36.739	106.153
LMSV	5-step	-0.004	0.054	0.032	-19.635	92.494
Ar1SV Spline	5-step	-0.009	0.054	0.032	-50.959	116.685
LMSV Spine	5-step	-0.007	0.056	0.033	-37.077	107.502
Ar1SV	22-step	-0.008	0.038	0.025	-26.610	76.615
LMSV	22-step	-0.003	0.038	0.026	-8.240	68.538
Ar1SV Spline	22-step	-0.014	0.041	0.028	-70.712	112.942
LMSV Spine	22-step	-0.008	0.400	0.027	-37.057	86.982

Table 6.14 – Out of Sample Forecast Measures: Litecoin and S&P500

	mean	sd	0.025quant	0.5quant	0.975quant
μ	-12.3	1.3	-14.9	-12.3	-9.81
H	0.984	0.00149	0.981	0.984	0.987
τ_h	0.212	0.018	0.179	0.211	0.25

Table 6.15 – Posterior Estimated Paramaters - LMSV for Bitcoin 10-minute returns

We also performed an out-of-sample forecasting experiment, comparing the predictive performance of five different specifications for forecast the realized volatility. In this experiment we left the last 22 days of the sample to perform predictions. In this experiment we compared the out-of-sample predictions using an ARFIMA model (0,d,0) applied directly for the measure of realized variance, LMSV model specifications using 10-minute returns data with and without the seasonal component (LMSV and LMSV Seas, respectively) and specifications of the SV model used a traditional AR(1) structure, also

with and without including the intraday seasonal component (Ar1SV and Ar1SV Seas). The estimation procedure is based on a growing window, using the entire sample up to the date before the forecasted day, and thus we perform the out-of-sample forecast for the last 22 observations of the sample.

	ME	RMSE	MAE
Arfima RVol	-0.00050851	0.0014185	0.0012083
LMSV	-0.0015462	0.0018967	0.001784
LMSV Seas	-0.0010842	0.0015719	0.0014445
Ar1SV	0.00022641	0.0010715	0.00071836
Ar1SV Seas	0.0002231	0.0010713	0.00072011

Table 6.16 – Forecast Measures for the Realized Variance

Table 6.16 shows the forecast results obtained in this experiment. We can observe that for this sample and sample period the Ar1SV and Ar1SV Seas models present the best predictive performances, followed by the direct prediction of an ARFIMA model (0,d,0) for the realized volatility, and then the models based on LMSV. Although in this experiment the LMSV model had a relatively lower performance, it is important to note that it is possible to obtain estimates for this model even with a large sample size and a simultaneous inclusion of the seasonal component for the intraday conditional variance process.

6.8 Conclusion

In this work we introduce an estimation methodology for long memory stochastic volatility models using an approximation of a fractional Gaussian Noise process as a Gaussian Markov Random Field, and from this approach we use the Integrated Nested Laplace Approximations methodology to perform the Bayesian estimation of parameters and latent variables.

This methodology is a useful addition to the set of tools used in the modeling of conditional volatility in time series, since it is computationally efficient and can be extended to add additional latent factors. In this work we show that a simple extension with the addition of a smooth variation component to model the variation in the long-run average of the conditional variance, analogous to a spline, allows gains in terms of in-sample fit and out-of-sample forecasting. Another extension introduced is that of an LMSV model with inclusion of seasonal patterns, which is especially important in the modeling of intraday returns.

The computational efficiency of this method is especially important in the modeling of high frequency data, characterized by a large number of observations. Sample size may

be an important limitation for other methods of conditional volatility estimation, such as quasi-maximum likelihood estimators and MCMC-based methods. We show how the proposed estimator can be used to construct measures of realized variance, and also make out-of-sample forecasts for this measure incorporating the entire dependence structure in the conditional variance observed in the intraday returns.

This methodology can be generalized in several directions. It is possible to construct multifactor models, for example with permanent and transient components or factors with short and long memory in the same model. It is possible to formulate a model with both Ar1 and fGn components, and to verify if there are different memory patterns in the conditional variance. A model with Ar1 and fGn components would be similar to an ARFIMA (1,d,0) specification for the conditional variance of the process, which may be a useful specification for some time series. Another class of possible extensions is the formulation of multivariate models of stochastic volatility with long memory, that can be constructed by increasing the observation state equations. We leave these possible extensions for other works.

6.9 Appendix - Specification Details

In all model the mean parameters (that is, μ in SV and LMSV models) have Normal prior with large known variance. Additional prior parameters are detailed below.

SV Model

$$r_t = \exp(h_t/2)\varepsilon_t, \quad \varepsilon_t \sim N(0, 1), \quad (6.27)$$

$$h_t = \mu + \phi(h_{t-1} - \mu) + z_t, \quad z_t \sim N(0, 1/(1 - \phi^2)\tau). \quad (6.28)$$

This model has three parameters: μ , ϕ , and τ .

The AR(1) structure has two hyperparameters θ_1 and θ_2

$$\theta_1 = \log(\kappa), \quad \kappa = \tau(1 - \phi^2), \quad (6.29)$$

$$\theta_2 = \log\left(\frac{1 + \phi}{1 - \phi}\right). \quad (6.30)$$

With prior distributions

$$\theta_1 \sim pc.prec(0.2, 0.008), \quad (6.31)$$

$$\theta_2 \sim pc.cor1(0.9, 0.9). \quad (6.32)$$

LMSV Model

$$r_t = \exp(h_t/2)\varepsilon_t, \quad \varepsilon_t \sim N(0, 1), \quad (6.33)$$

$$h_t = \mu + fGn(H, \tau_h), \quad (6.34)$$

There are three estimated parameters: μ , H , and τ_h .

The fGn process h_t has two hyperparameters θ_1 and θ_2 .

$$\tau = \exp(\theta_1), \quad \theta_1 \sim pc.prec(0.2, 0.008) \quad (6.35)$$

$$H = \frac{1}{2} \left(\frac{\exp(\theta_2)}{1 + \exp(\theta_2)} \right), \quad \theta_2 \sim pc.cor1(0.9, 0.9) \quad (6.36)$$

SV Spline Model

In the SV Spline model log volatility follows

$$r_t = \exp(h_t + s_t)/2\varepsilon_t, \quad \varepsilon_t \sim N(0, 1), \quad (6.37)$$

$$h_t = \phi h_{t-1} + \eta_t, \quad h_t \sim N(0, 1/\tau_h), \quad (6.38)$$

$$\Delta^2 s_t = s_t - 2s_{t+1} + s_{t+2}, \quad s_t \sim N(0, 1/\tau_s) \quad (6.39)$$

The second-order random walk spline component has one hyperparameters θ_1 which is the marginal precision of the component;

$$\theta_1 = \log(\tau_s), \quad (6.40)$$

$$\theta_1 \sim logGamma(0, 0.15). \quad (6.41)$$

This model has three parameters τ_s , ϕ , τ_h .

The AR(1) term has prior structure similar to the standard SV Model, with two hyperparameters θ_3, θ_2 .

$$\theta_2 = \log(\kappa), \quad \kappa = (1 - \phi), \quad (6.42)$$

$$\theta_3 = \log \left(\frac{1 + \phi}{1 - \phi} \right) \quad (6.43)$$

$$\theta_2 \sim pc.prec(0.2, 0.008), \quad (6.44)$$

$$\theta_3 \sim pc.cor1(0.9, 0.9). \quad (6.45)$$

Spline LMSV Model

In the Spline LMSV model log volatility follows

$$r_t = \exp((h_t + s_t)/2)\varepsilon_t, \quad \varepsilon_t \sim N(0, 1), \quad (6.46)$$

$$h_t = \mu + fGn(H, \tau), \quad (6.47)$$

$$\Delta^2 s_t = s_t - 2s_{t+1} + s_{t+2}, \quad s_t \sim N(0, 1/\tau_s). \quad (6.48)$$

The second-order random walk spline component has one hyperparameters θ_1 which is the precision of the i.i.d Normal innovations ν_t .

$$\theta_1 = \log(\tau_s), \quad (6.49)$$

$$\theta_1 \sim \log\text{Gamma}(0, 0.15). \quad (6.50)$$

The fGn process y_t has two hyperparameters θ_2 and θ_3 .

$$\tau = \exp(\theta_2), \quad \theta_2 \sim \text{pc.prec}(0.2, 8e - 03) \quad (6.51)$$

$$H = \frac{1}{2} \left(\frac{\exp(\theta_3)}{1 + \exp(\theta_2)} \right), \quad \theta_3 \sim \text{pc.cor1}(0.9, 0.9) \quad (6.52)$$

Seasonal LMSV Model

$$h_t = \text{seas}_t + y_t + \eta_t, \quad \eta_t \sim N(0, \tau_h) \quad (6.53)$$

$$\text{seas}_t = \text{seas}_t + 1 + \dots + \text{seas}_{t+m+1} + \nu_t, \text{seas}_t \sim N(0, \tau_{\text{seas}}) \quad (6.54)$$

This model has three estimated parameters τ_{seas} , τ_h , and H . Periodicity m is 144, which reflects daily frequency with returns 10 minutes apart.

The seasonal component seas_t has one hyperparameter θ_1

$$\theta_1 = \log(\tau_{\text{seas}}) \quad (6.55)$$

$$\theta_1 \sim \log\text{Gamma}(0, 0.15). \quad (6.56)$$

The fGn component has two hyperparameters θ_2 , θ_3

$$\theta_1 = \log(\kappa), \quad k = \tau(1 - \rho^2), \quad (6.57)$$

$$\theta_2 = \log \left(\frac{1 + \rho}{1 - \rho} \right) \quad (6.58)$$

We employ penalized complexity priors introduced by Sørbye and Rue (2018) for both precision τ and persistence ρ .

Bibliography

- AGUILAR, O.; WEST, M. Bayesian dynamic factor models and portfolio allocation. *Journal of Business & Economic Statistics*, Taylor & Francis, v. 18, n. 3, p. 338–357, 2000.
- AIT-SAHALIA, Y.; KIMMEL, R. Maximum likelihood estimation of stochastic volatility models. *Journal of Financial Economics*, Elsevier, v. 83, n. 2, p. 413–452, 2007.
- AKSAMIT, A.; JEANBLANC, M. *Enlargement of Filtration with Finance in View*. [S.l.]: Springer Briefs in Mathematics, 2017.
- AL-YAHYAEI, K. H.; MENSI, W.; YOON, S.-M. Efficiency, multifractality, and the long-memory property of the bitcoin market: A comparative analysis with stock, currency, and gold markets. *Finance Research Letters*, v. 27, p. 228 – 234, 2018.
- ALBERT, J. H.; CHIB, S. Bayesian analysis of binary and polychotomous response data. *Journal of the American statistical Association*, Taylor & Francis, v. 88, n. 422, p. 669–679, 1993.
- ALVAREZ-RAMIREZ, J.; RODRIGUEZ, E.; IBARRA-VALDEZ, C. Long-range correlations and asymmetry in the bitcoin market. *Physica A: Statistical Mechanics and its Applications*, v. 492, p. 948 – 955, 2018.
- ANDERSEN, L. B.; PITERBARG, V. V. Moment explosions in stochastic volatility models. *Finance and Stochastics*, Springer, v. 11, n. 1, p. 29–50, 2007.
- ANDERSEN, T. G.; CHUNG, H.-J.; SØRENSEN, B. E. Efficient method of moments estimation of a stochastic volatility model: A Monte Carlo study. *Journal of Econometrics*, v. 91, n. 1, p. 61 – 87, 1999. ISSN 0304-4076.
- ANDERSEN, T. G.; SØRENSEN, B. E. Gmm estimation of a stochastic volatility model: A monte carlo study. *Journal of Business & Economic Statistics*, Taylor Francis, v. 14, n. 3, p. 328–352, 1996.
- ARDIA, D.; BLUTEAU, K.; RÜEDE, M. Regime changes in bitcoin garch volatility dynamics. *Finance Research Letters*, 2018.
- ARTECHE, J. Gaussian semiparametric estimation in long memory in stochastic volatility and signal plus noise models. *Journal of Econometrics*, v. 119, p. 119–131, 2004.
- ASAI, M. Comparison of MCMC methods for estimating stochastic volatility models. *Computational Economics*, v. 25, n. 3, p. 281–301, Jun 2005.
- BALCILAR, M. et al. Can volume predict bitcoin returns and volatility? a quantiles-based approach. *Economic Modelling*, Elsevier, v. 64, p. 74–81, 2017.
- BALDEAUX, J.; IGNATIEVA, K.; PLATEN, E. Detecting money market bubbles. *Journal of Banking & Finance*, Elsevier, v. 87, p. 369–379, 2018.

- BARIVIERA, A. F. The inefficiency of bitcoin revisited: A dynamic approach. *Economics Letters*, Elsevier, v. 161, p. 1–4, 2017.
- BARIVIERA, A. F. et al. Some stylized facts of the bitcoin market. *Physica A: Statistical Mechanics and its Applications*, Elsevier, v. 484, p. 82–90, 2017.
- BASAK, G. K.; CHAN, N. H.; PALMA, W. On the approximation of long memory processes by an ARMA model. *Journal of Forecasting*, v. 20, p. 367–389, 2001.
- BAUR, D. G.; DIMPFL, T. Asymmetric volatility in cryptocurrencies. *Economics Letters*, Elsevier, v. 173, p. 148–151, 2018.
- BAUR, D. G.; DIMPFL, T.; KUCK, K. Bitcoin, gold and the us dollar—a replication and extension. *Finance Research Letters*, Elsevier, v. 25, p. 103–110, 2018.
- BERAN, J. *Statistics for long-memory processes*. [S.l.]: Routledge, 2017.
- BERAN, J.; SCHÜTZNER, M.; GHOSH, S. From short to long memory: Aggregation and estimation. *Computational Statistics & Data Analysis*, Elsevier, v. 54, n. 11, p. 2432–2442, 2010.
- BETANCOURT, M. A conceptual introduction to Hamiltonian Monte Carlo. *arXiv preprint arXiv:1701.02434*, 2017.
- BLOOMBERG. *Jamie Dimon Slams Bitcoin as a ‘Fraud’*. 2017. Disponível em: <<https://www.bloomberg.com/news/articles/2017-09-12/jpmorgan-s-ceo-says-he-d-fire-traders-who-bet-on-fraud-bitcoin>>.
- BOLLERSLEV, T. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, v. 31, n. 3, p. 307 – 327, 1986.
- BOURI, E. et al. Modelling long memory volatility in the Bitcoin market: Evidence of persistence and structural breaks. *International Journal of Finance & Economics*, v. 24, n. 1, p. 412–426, 2019.
- BOURI, E. et al. Bitcoin for energy commodities before and after the december 2013 crash: diversifier, hedge or safe haven? *Applied Economics*, Taylor & Francis, v. 49, n. 50, p. 5063–5073, 2017.
- BOURI, E. et al. On the hedge and safe haven properties of bitcoin: Is it really more than a diversifier? *Finance Research Letters*, Elsevier, v. 20, p. 192–198, 2017.
- BOURI, E.; SHAHZAD, S. J. H.; ROUBAUD, D. Co-explosivity in the cryptocurrency market. *Finance Research Letters*, Elsevier, forthcoming, 2018.
- BREIDT, F. J.; CRATO, N.; LIMA, P. D. The detection and estimation of long memory in stochastic volatility. *Journal of Econometrics*, Elsevier, v. 83, n. 1-2, p. 325–348, 1998.
- CAMERER, C. Bubbles and fads in asset prices. *Journal of Economic Surveys*, Wiley Online Library, v. 3, n. 1, p. 3–41, 1989.
- CAMERER, C.; WEIGELT, K. Information mirages in experimental asset markets. *Journal of Business*, JSTOR, v. 64, n. 4, p. 463–493, 1991.

- CHAIM, P.; LAURINI, M. P. Volatility and return jumps in bitcoin. *Economics Letters*, Elsevier, v. 173, p. 158–163, 2018.
- CHAIM, P.; LAURINI, M. P. Is bitcoin a bubble? *Physica A: Statistical Mechanics and its Applications*, Elsevier, v. 517, p. 222 – 232, 2019.
- CHAIM, P.; LAURINI, M. P. Nonlinear dependence in cryptocurrency markets. *The North American Journal of Economics and Finance*, v. 48, p. 32 – 47, 2019.
- CHAN, N. H.; PALMA, W. State space modelling of long memory processes. *Annals of Statistics*, v. 26, p. 719–740, 1998.
- CHAN, N. H.; PETRIS, G. Long memory stochastic volatility: A Bayesian approach. *Communication in Statistics – Theory and Methods*, v. 29, p. 1367—1378, 2000.
- CHARFEDDINE, L.; MAOUCI, Y. Are shocks on the returns and volatility of cryptocurrencies really persistent? *Finance Research Letters*, Elsevier, 2018.
- CHARLES, A.; DARNÉ, O. Volatility estimation for bitcoin: Replication and robustness. *International Economics*, Elsevier, v. 157, p. 23–32, 2019.
- CHARLES, A.; DARNÉ, O.; KIM, J. H. Small sample properties of alternative tests for martingale difference hypothesis. *Economics Letters*, Elsevier, v. 110, n. 2, p. 151–154, 2011.
- CHEAH, E.-T.; FRY, J. Speculative bubbles in bitcoin markets? an empirical investigation into the fundamental value of bitcoin. *Economics Letters*, Elsevier, v. 130, p. 32–36, 2015.
- CHEAH, E.-T. et al. Long memory interdependency and inefficiency in bitcoin markets. *Economics Letters*, v. 167, p. 18 – 25, 2018. ISSN 0165-1765.
- CHEUNG, A.; ROCA, E.; SU, J.-J. Crypto-currency bubbles: an application of the Phillips–Shi–Yu (2013) methodology on Mt. Gox Bitcoin prices. *Applied Economics*, Taylor & Francis, v. 47, n. 23, p. 2348–2358, 2015.
- CHIB, S.; NARDARI, F.; SHEPHARD, N. Markov chain Monte Carlo methods for stochastic volatility models. *Journal of Econometrics*, v. 108, n. 2, p. 281 – 316, 2002.
- CHIB, S.; NARDARI, F.; SHEPHARD, N. Analysis of high dimensional multivariate stochastic volatility models. *Journal of Econometrics*, Elsevier, v. 134, n. 2, p. 341–371, 2006.
- CHRISTENSEN, B. J.; NIELSEN, M. O. The effect of long memory in volatility on stock market fluctuations. *The Review of Economics and Statistics*, v. 89, n. 4, p. 684–700, 2007.
- CHRISTOFFERSEN, P. Evaluating interval forecasts. *International Economic Review*, v. 39, p. 841–842, 1998.
- CIAIAN, P.; RAJCANIOVA, M.; KANCS, d. The economics of bitcoin price formation. *Applied Economics*, Taylor & Francis, v. 48, n. 19, p. 1799–1815, 2016.

CLARK, T. E.; WEST, K. D. Approximately normal tests for equal predictive accuracy in nested models. *Journal of econometrics*, Elsevier, v. 138, n. 1, p. 291–311, 2007.

COCHRANE, J. H. *Stocks as money: convenience yield and the tech-stock bubble*. [S.l.], 2002.

CORBET, S.; LUCEY, B.; YAROVAYA, L. Datestamping the bitcoin and ethereum bubbles. *Finance Research Letters*, Elsevier, v. 26, p. 81–88, 2018.

CORCUERA, J. M.; VADIVIA, A. Enlargements of filtrations and applications. *arXiv preprint arXiv:1201.5870*, 2012.

CRATO, N.; RAY, B. K. Semi-parametric smoothing estimators for long memory processes with added noise. *Journal of Statistical Planning and Inference*, v. 105, p. 283–297, 2002.

CREAL, D.; KOOPMAN, S. J.; LUCAS, A. Generalized autoregressive score models with applications. *Journal of Applied Econometrics*, v. 28, n. 5, p. 777–795, 2013.

DANDAPANI, A.; PROTTER, P. Strict local martingales via filtration enlargement. *arXiv preprint arXiv:1608.06361*, 2016.

DANDAPANI, A.; PROTTER, P. Strict local martingales and the khasminskii test for explosions. *Stochastic Processes and their Applications*, 2019. ISSN 0304-4149. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0304414918307518>>.

DELBAEN, F.; SCHACHERMAYER, W. The fundamental theorem of asset pricing for unbounded stochastic processes. *Mathematische annalen*, Springer, v. 312, n. 2, p. 215–250, 1998.

DELBAEN, F.; SCHACHERMAYER, W. *The mathematics of arbitrage*. [S.l.]: Birkhäuser, 2006.

DEO, R.; HURVICH, C.; LU, Y. Forecasting realized volatility using a long memory stochastic volatility model: Estimation, prediction and seasonal adjustment. *Journal of Econometrics*, v. 131, p. 29–58, 2006.

DIEBOLD, F. X.; INOUE, A. Long memory and regime switching. *Journal of econometrics*, Elsevier, v. 105, n. 1, p. 131–159, 2001.

DIEBOLD, F. X.; MARIANO, R. S. Comparing predictive accuracy. *Journal of Business & economic statistics*, Taylor & Francis, v. 20, n. 1, p. 134–144, 2002.

DOMÍNGUEZ, M. A.; LOBATO, I. N. Testing the martingale difference hypothesis. *Econometric Reviews*, Taylor & Francis, v. 22, n. 4, p. 351–377, 2003.

DUANE, S. et al. Hybrid monte carlo. *Physics Letters B*, Elsevier, v. 195, n. 2, p. 216–222, 1987.

DYHRBERG, A. H. Bitcoin, gold and the dollar—a garch volatility analysis. *Finance Research Letters*, Elsevier, v. 16, p. 85–92, 2016.

DYHRBERG, A. H. Hedging capabilities of bitcoin. is it the virtual gold? *Finance Research Letters*, Elsevier, v. 16, p. 139–144, 2016.

- EHLERS, R.; ZEVALLOS, M. Bayesian estimation and prediction of stochastic volatility models via inla. *Communications in Statistics - Simulation and Computation*, Taylor Francis, v. 44, n. 3, p. 683–693, 2015.
- ENGLE, R. F. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, [Wiley, Econometric Society], v. 50, n. 4, p. 987–1007, 1982.
- ENGLE, R. F.; RANGEL, J. G. The spline-GARCH model for low-frequency volatility and its global macroeconomic causes. *Review of Financial Studies*, v. 21, n. 3, p. 1187–1222, 2008.
- FERRAZ, R. O.; HOTTA, L. K. Quasi-maximum likelihood estimation of long-memory stochastic volatility models. *Brazilian Review of Econometrics*, Springer, v. 27, n. 2, p. 225–233, 2007.
- FLEMING, J.; KIRBY, C. Long memory in volatility and trading volume. *Journal of Banking & Finance*, v. 35, n. 7, p. 1714 – 1726, 2011.
- FLOOD, R. P.; HODRICK, R. J. Asset price volatility, bubbles, and process switching. *The Journal of Finance*, Wiley Online Library, v. 41, n. 4, p. 831–842, 1986.
- FLORENS-ZMIROU, D. On estimating the diffusion coefficient from discrete observations. *Journal of Applied Probability*, Cambridge University Press, v. 30, n. 4, p. 790–804, 1993.
- FRY, J. Booms, busts and heavy-tails: The story of bitcoin and cryptocurrency markets? *Economics Letters*, v. 171, p. 225 – 229, 2018. ISSN 0165-1765.
- GAMERMAN, D.; LOPES, H. F. *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference*. [S.l.]: Chapman & Hall, 2006.
- GANDAL, N. et al. Price manipulation in the bitcoin ecosystem. *Journal of Monetary Economics*, Elsevier, v. 95, p. 86–96, 2018.
- GARCIA, D. et al. The digital traces of bubbles: feedback cycles between socio-economic signals in the bitcoin economy. *Journal of the Royal Society Interface*, The Royal Society, v. 11, n. 99, p. 20140623, 2014.
- GEWEKE, J.; PORTER-HUDAK, S. The estimation and application of long memory time series models. *Journal of time series analysis*, Wiley Online Library, v. 4, n. 4, p. 221–238, 1983.
- GEWEKE, J.; ZHOU, G. Measuring the pricing error of the arbitrage pricing theory. *The review of financial studies*, Oxford University Press, v. 9, n. 2, p. 557–587, 1996.
- GKILLAS, K.; KATSIAMPA, P. An application of extreme value theory to cryptocurrencies. *Economics Letters*, v. 164, p. 109 – 111, 2018. ISSN 0165-1765.
- GRANGER, C. W. Long memory relationships and the aggregation of dynamic models. *Journal of econometrics*, Elsevier, v. 14, n. 2, p. 227–238, 1980.
- GRONWALD, M. *The Economics of Bitcoins - Market Characteristics and Price Jumps*. [S.l.], 2014.

- HALDRUP, N.; VALDÉS, J. E. V. Long memory, fractional integration, and cross-sectional aggregation. *Journal of econometrics*, Elsevier, v. 199, n. 1, p. 1–11, 2017.
- HAN, Y. Asset allocation with a high dimensional latent factor stochastic volatility model. *The Review of Financial Studies*, Oxford University Press, v. 19, n. 1, p. 237–271, 2005.
- HARVEY, A. *Forecasting, Structural Time Series and the Kalman Filter*. [S.l.]: Cambridge University Press, 1989.
- HARVEY, A.; RUIZ, E.; SHEPHARD, N. G. *Multivariate stochastic variance model*. [S.l.], 1992. Financial Markets Group discussion paper (London School of Economics. London).
- HARVEY, A.; TRIMBUR, T. Trend estimation and the Hodrick-Prescott filter. *Journal of the Japan Statistical Society*, v. 38, n. 1, p. 41–49, 2008.
- HARVEY, A. C. Stochastic volatility models with long memory. In: KNIGHT, J.; SATCHELL, E. (Ed.). *Forecasting Volatility in Financial Markets*. [S.l.]: Butterworth-Haineman, London, 1998. p. 307–320.
- HARVEY, A. C.; JAEGER, A. Detrending, stylized facts and the business cycle. *Journal of Applied Econometrics*, v. 8, n. 3, p. 231–247, 1993.
- HESTON, S. L. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies*, Society of Financial Studies, v. 6, n. 2, p. 327–343, 1993.
- HILLEBRAND, E.; MEDEIROS, M. C. Nonlinearity, breaks, and long-range dependence in time-series models. *Journal of Business & Economic Statistics*, Taylor & Francis, v. 34, n. 1, p. 23–41, 2016.
- HOSKING, J. R. M. Fractional differencing. *Biometrika*, v. 68, n. 1, p. 165–176, 1981.
- HURVICH, C. M.; SIMONOFF, J. S.; TSAI, C.-L. Smoothing parameter selection in nonparametric regression using an improved akaike information criterion. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, Wiley Online Library, v. 60, n. 2, p. 271–293, 1998.
- HURVICH, C. M.; SOULIER, P. Stochastic volatility models with long memory. In: *Handbook of Financial Time Series*. [S.l.]: Springer, 2009. p. 345–354.
- JACOD, J. Grossissement initial, hypothese (h) et theoreme de girsanov. In: *Grossissements de filtrations: exemples et applications*. [S.l.]: Springer, 1985. p. 15–35.
- JARROW, R. Detecting asset price bubbles. *Journal of Derivatives*, Euromoney Institutional Investor PLC, v. 20, n. 1, p. 30–34, 2012.
- JARROW, R. Testing for asset price bubbles: three new approaches. *Quantitative Finance Letters*, Taylor & Francis, v. 4, n. 1, p. 4–9, 2016.
- JARROW, R.; KCHIA, Y.; PROTTER, P. How to detect an asset bubble. *SIAM Journal on Financial Mathematics*, SIAM, v. 2, n. 1, p. 839–865, 2011.

- JARROW, R.; KCHIA, Y.; PROTTER, P. Is there a bubble in LinkedIn's stock price? *arXiv preprint arXiv:1105.5717*, 2011.
- JARROW, R.; PROTTER, P.; SHIMBO, K. In: FU, M. C. et al. (Ed.). *Advances in Mathematical Finance. Applied and Numerical Harmonic Analysis*. [S.l.]: Birkhauser Boston, 2007. cap. Asset Price Bubbles in Complete Markets.
- JARROW, R. A.; PROTTER, P.; SHIMBO, K. Asset price bubbles in incomplete markets. *Mathematical Finance*, Wiley Online Library, v. 20, n. 2, p. 145–185, 2010.
- JEULIN, T. Grossissement d'une filtration et applications. *Séminaire de Probabilités*, v. 15, p. 574–609, 1978.
- JEULIN, T.; YOR, M. Grossissement d'une filtration et semi-martingales: formules explicites. *Séminaire de Probabilités*, v. 12, p. 78–97, 1978.
- JEULIN, T.; YOR, M. *Grossissements de filtrations: exemples et applications, Lecture Notes in Mathematics 1118*. [S.l.]: Springer, 1985.
- JIANG, G. J.; KNIGHT, J. L. A nonparametric approach to the estimation of diffusion processes, with an application to a short-term interest rate model. *Econometric Theory*, Cambridge University Press, v. 13, n. 5, p. 615–645, 1997.
- JIANG, Y.; NIE, H.; RUAN, W. Time-varying long-term memory in bitcoin market. *Finance Research Letters*, v. 25, p. 280 – 284, 2018. ISSN 1544-6123.
- JOHANES, M.; POLSON, N. Handbook of financial econometrics. In: _____. [S.l.]: Elsevier-North-Holland, 2005. cap. MCMC for Financial Econometrics.
- KASTNER, G. Dealing with stochastic volatility in time series using the R package stochvol. *Journal of Statistical Software*, v. 69, n. 5, p. 1–30, 2016.
- KASTNER, G. Sparse Bayesian time-varying covariance estimation in many dimensions. *Journal of Econometrics*, v. 210, n. 1, p. 98 – 115, 2019. Annals Issue in Honor of John Geweke “Complexity and Big Data in Economics and Finance: Recent Developments from a Bayesian Perspective”.
- KASTNER, G.; FRÜHWIRTH-SCHNATTER, S. Ancillarity-sufficiency interweaving strategy (asis) for boosting MCMC estimation of stochastic volatility models. *Comput. Stat. Data Anal.*, Elsevier Science Publishers B. V., Amsterdam, The Netherlands, The Netherlands, v. 76, n. C, p. 408–423, ago. 2014. ISSN 0167-9473.
- KASTNER, G.; FRÜHWIRTH-SCHNATTER, S. Ancillarity-sufficiency interweaving strategy (asis) for boosting mcmc estimation of stochastic volatility models. *Computational Statistics & Data Analysis*, v. 76, p. 408 – 423, 2014. CFEnetwork: The Annals of Computational and Financial Econometrics.
- KATSIAMPA, P. Volatility estimation for bitcoin: A comparison of garch models. *Economics Letters*, Elsevier, v. 158, p. 3–6, 2017.
- KATSIAMPA, P. Volatility co-movement between bitcoin and ether. *Finance Research Letters*, 2018. ISSN 1544-6123.

- KIM, S.; SHEPHARD, N.; CHIB, S. Stochastic volatility: Likelihood inference and comparison with ARCH models. *The Review of Economic Studies*, [Oxford University Press, Review of Economic Studies, Ltd.], v. 65, n. 3, p. 361–393, 1998.
- KOTANI, S. On a condition that one-dimensional diffusion processes are martingales. In: *In Memoriam Paul-André Meyer*. [S.l.]: Springer, 2006. p. 149–156.
- KOUTMOS, D. Return and volatility spillovers among cryptocurrencies. *Economics Letters*, v. 173, p. 122 – 127, 2018. ISSN 0165-1765.
- KRISTOUFEK, L. Bitcoin meets google trends and wikipedia: Quantifying the relationship between phenomena of the internet era. *Scientific reports*, Nature Publishing Group, v. 3, p. 3415, 2013.
- KRISTOUFEK, L. Bitcoin meets google trends and wikipedia: Quantifying the relationship between phenomena of the internet era. *Scientific Reports*, v. 3, n. 3415, 2013.
- KRISTOUFEK, L. What are the main drivers of the bitcoin price? evidence from wavelet coherence analysis. *PloS one*, Public Library of Science, v. 10, n. 4, p. e0123923, 2015.
- KRISTOUFEK, L. On bitcoin markets (in)efficiency and its evolution. *Physica A: Statistical Mechanics and its Applications*, v. 503, p. 257 – 262, 2018. ISSN 0378-4371.
- KRUGMAN, P. *Bubble, Bubble, Fraud and Trouble*. 2018. Disponível em: <<https://www.nytimes.com/2018/01/29/opinion/bitcoin-bubble-fraud.html>>.
- KWIATKOWSKI, D. et al. Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? *Journal of econometrics*, Elsevier, v. 54, n. 1-3, p. 159–178, 1992.
- LAHMIRI, S.; BEKIROU, S.; SALVI, A. Long-range memory, distributional variation and randomness of bitcoin volatility. *Chaos, Solitons & Fractals*, Elsevier, v. 107, p. 43–48, 2018.
- LAI, T. L.; XING, H. Structural change as an alternative to long memory in financial time series. In: *Econometric Analysis of Financial and Economic Time Series*. [S.l.]: Emerald Group Publishing Limited, 2006. p. 205–224.
- LAURINI, M.; MAUAD, R. The stochastic volatility model with random jumps and its application to brl/usd exchange rate. *Economics Bulletin*, v. 34, n. 2, p. 1002–1011, 2014.
- LAURINI, M. P.; HOTTA, L. K. GMC/GEL estimation of stochastic volatility models. *Communications in Statistics - Simulation and Computation*, Taylor Francis, v. 46, n. 9, p. 6828–6844, 2017.
- LAURINI, M. P.; MAUAD, R.; AUIBE, F. L. *Multivariate Stochastic Volatility-Double Jump Model: An Application for Oil Assets*. [S.l.], 2016.
- LIEBERMAN, O.; PHILLIPS, P. C. B. Refined inference on long memory in realized volatility. *Econometric Reviews*, Taylor Francis, v. 27, n. 1-3, p. 254–267, 2008.
- LINDGREN, F.; RUE, H. On the second-order random walk model for irregular locations. *Scandinavian Journal of Statistics*, v. 35, n. 4, p. 691–700, 2008.

- LIONS, P.-L.; MUSIELA, M. Correlations and bounds for stochastic volatility models. In: ELSEVIER. *Annales de l'Institut Henri Poincaré (C) Non Linear Analysis*. [S.l.], 2007. v. 24, n. 1, p. 1–16.
- LOPES, H. F.; CARVALHO, C. M. Factor stochastic volatility with time varying loadings and markov switching regimes. *Journal of Statistical Planning and Inference*, Elsevier, v. 137, n. 10, p. 3082–3091, 2007.
- MAASOUMI, E.; MCALEER, M. Realized volatility and long memory: An overview. *Econometric Reviews*, Taylor Francis, v. 27, n. 1-3, p. 1–9, 2008.
- MACDONELL, A. Popping the bitcoin bubble: An application of log-periodic power law modeling to digital currency. 2014.
- MANDELBROT, B. B.; NESS, J. W. V. Fractional Brownian motions, fractional noises and applications. *SIAM review*, SIAM, v. 10, n. 4, p. 422–437, 1968.
- MARIANI, L. A.; LAURINI, M. P. Implicit inflation and risk premiums in the brazilian fixed income market. *Emerging Markets Finance and Trade*, Taylor & Francis, n. just-accepted, 2016.
- MARTINO, S. et al. Estimating stochastic volatility models using integrated nested Laplace approximations. *The European Journal of Finance*, Routledge, v. 17, n. 7, p. 487–503, 2011.
- MCALEER, M.; MEDEIROS, M. C. Realized volatility: A review. *Econometric Reviews*, Taylor Francis, v. 27, n. 1-3, p. 10–45, 2008.
- MENSI, W.; AL-YAHYAE, K. H.; KANG, S. H. Structural breaks and double long memory of cryptocurrency prices: A comparative analysis from Bitcoin and Ethereum. *Finance Research Letters*, 2018.
- MIJATOVIĆ, A.; URUSOV, M. On the martingale property of certain local martingales. *Probability Theory and Related Fields*, Springer, v. 152, n. 1-2, p. 1–30, 2012.
- MIKHAILOV, S.; NÖGEL, U. *Heston's stochastic volatility model: Implementation, calibration and some extensions*. [S.l.]: John Wiley and Sons, 2004.
- NELSON, D. B. *The time series behavior of stock market volatility and returns*. [S.l.], 1988. Unpublished Ph.D.dissertation (Massachusetts Institute of Technology, Cambridge, MA).
- OBAYASHI, Y.; PROTTER, P.; YANG, S. The lifetime of a financial bubble. *Mathematics and Financial Economics*, Springer, v. 11, n. 1, p. 45–62, 2017.
- PAIGE, R. L.; TRINDADE, A. A. The Hodrick-Prescott Filter: A special case of penalized spline smoothing. *Electronic Journal of Statistics*, The Institute of Mathematical Statistics and the Bernoulli Society, v. 4, p. 856–874, 2010.
- PEREZ, A.; RUIZ, E. Finite sample properties of a QML estimator of stochastic volatility models with long memory. *Economics Letters*, v. 70, p. 157–164, 2001.

PERRON, P.; QU, Z. Long-memory and level shifts in the volatility of stock market return indices. *Journal of Business & Economic Statistics*, Taylor & Francis, v. 28, n. 2, p. 275–290, 2010.

PHILLIP, A.; CHAN, J.; PEIRIS, S. On long memory effects in the volatility measure of cryptocurrencies. *Finance Research Letters*, 2018. ISSN 1544-6123.

PHILLIP, A.; CHAN, J.; PEIRIS, S. On long memory effects in the volatility measure of cryptocurrencies. *Finance Research Letters*, v. 28, p. 95 – 100, 2019.

PHILLIPS, P. C.; SHI, S.; YU, J. Testing for multiple bubbles: Historical episodes of exuberance and collapse in the S&P 500. *International Economic Review*, Wiley Online Library, v. 56, n. 4, p. 1043–1078, 2015.

PIIROINEN, P. et al. Asset price bubbles: An option-based indicator. *arXiv preprint arXiv:1805.07403*, 2018.

PROTTER, P. A mathematical theory of financial bubbles. In: *Paris-Princeton Lectures on Mathematical Finance 2013*. [S.l.]: Springer, 2013. p. 1–108.

QU, Z.; PERRON, P. A stochastic volatility model with random level shifts and its applications to s&p 500 and nasdaq return indices. *The Econometrics Journal*, Wiley Online Library, v. 16, n. 3, p. 309–339, 2013.

RUE, H.; MARTINO, S.; CHOPIN, N. Approximate Bayesian inference for latent gaussian models by using integrated nested Laplace approximations. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, Wiley Online Library, v. 71, n. 2, p. 319–392, 2009.

SANDMANN, G.; KOOPMAN, S. J. Estimation of stochastic volatility models via monte carlo maximum likelihood. *Journal of Econometrics*, v. 87, n. 2, p. 271 – 301, 1998. ISSN 0304-4076.

SCAILLET, O.; TRECCANI, A.; TREVISAN, C. High-frequency jump analysis of the bitcoin market. *Swiss Finance Institute Research Paper No. 17-19.*, 2017.

SCHARTH, M.; MEDEIROS, M. C. Asymmetric effects and long memory in the volatility of dow jones stocks. *International Journal of Forecasting*, Elsevier, v. 25, n. 2, p. 304–327, 2009.

SHACKLETON, M. B.; PONG, S.; TAYLOR, S. J. Distinguishing short and long memory volatility specifications. *The Econometrics Journal*, v. 11, n. 3, p. 617–637, 2008.

SIMPSON, D. et al. Penalising model component complexity: A principled, practical approach to constructing priors. *Statistical Science*, The Institute of Mathematical Statistics, v. 32, n. 1, p. 1–28, 02 2017.

SMITH, V. L.; SUCHANEK, G. L.; WILLIAMS, A. W. Bubbles, crashes, and endogenous expectations in experimental spot asset markets. *Econometrica*, JSTOR, v. 56, n. 5, p. 1119–1151, 1988.

SØRBYE, S. H.; MYRVOLL-NILSEN, E.; RUE, H. An approximate fractional Gaussian noise model with $\mathcal{O}(n)$ computational cost. *Statistics and Computing*, Springer, p. 1–13, 2017.

- SØRBYE, S. H.; RUE, H. Fractional Gaussian noise: Prior specification and model comparison. *Environmetrics*, v. 29, n. 5-6, 2018.
- SU, C.-W. et al. Testing for multiple bubbles in bitcoin markets: A generalized sup ADF test. *Japan and the World Economy*, Elsevier, v. 46, p. 56–63, 2018.
- TAQQU, M. S. Fractional Brownian motion and long-range dependence. *Theory and applications of long-range dependence*, Boston, MA: Birkhauser, p. 5–38, 2003.
- TAYLOR, S. J. *Modelling Financial Time Series*. [S.l.]: John Wiley& Sons, 1986.
- TIWARI, A. K. et al. Informational efficiency of bitcoin—an extension. *Economics Letters*, Elsevier, v. 163, p. 106–109, 2018.
- TROSTER, V. et al. Bitcoin returns and risk: A general garch and gas analysis. *Finance Research Letters*, 2018.
- URQUHART, A. The inefficiency of bitcoin. *Economics Letters*, Elsevier, v. 148, p. 80–82, 2016.
- VEHTARI, A.; GELMAN, A.; GABRY, J. Practical bayesian model evaluation using leave-one-out cross-validation and waic. *Statistics and Computing*, Springer, p. 1–20, 2016.
- WEI, W. C. Liquidity and market efficiency in cryptocurrencies. *Economics Letters*, Elsevier, v. 168, p. 21–24, 2018.
- YELOWITZ, A.; WILSON, M. Characteristics of bitcoin users: an analysis of google search data. *Applied Economics Letters*, Taylor & Francis, v. 22, n. 13, p. 1030–1036, 2015.
- ZARGAR, F. N.; KUMAR, D. Long range dependence in the bitcoin market: A study based on high-frequency data. *Physica A: Statistical Mechanics and its Applications*, v. 515, p. 625 – 640, 2019.
- ZHANG, L.; MYKLAND P. A. AIT-SAHALIA, Y. A tale of two time scales: Determining integrated volatility with noisy high-frequency data. *Journal of the American Statistical Association*, v. 100, p. 1394–1411, 2005.
- ZHOU, X.; NAKAJIMA, J.; WEST, M. Bayesian forecasting and portfolio decisions using dynamic dependent sparse factor models. *International Journal of Forecasting*, Elsevier, v. 30, n. 4, p. 963–980, 2014.