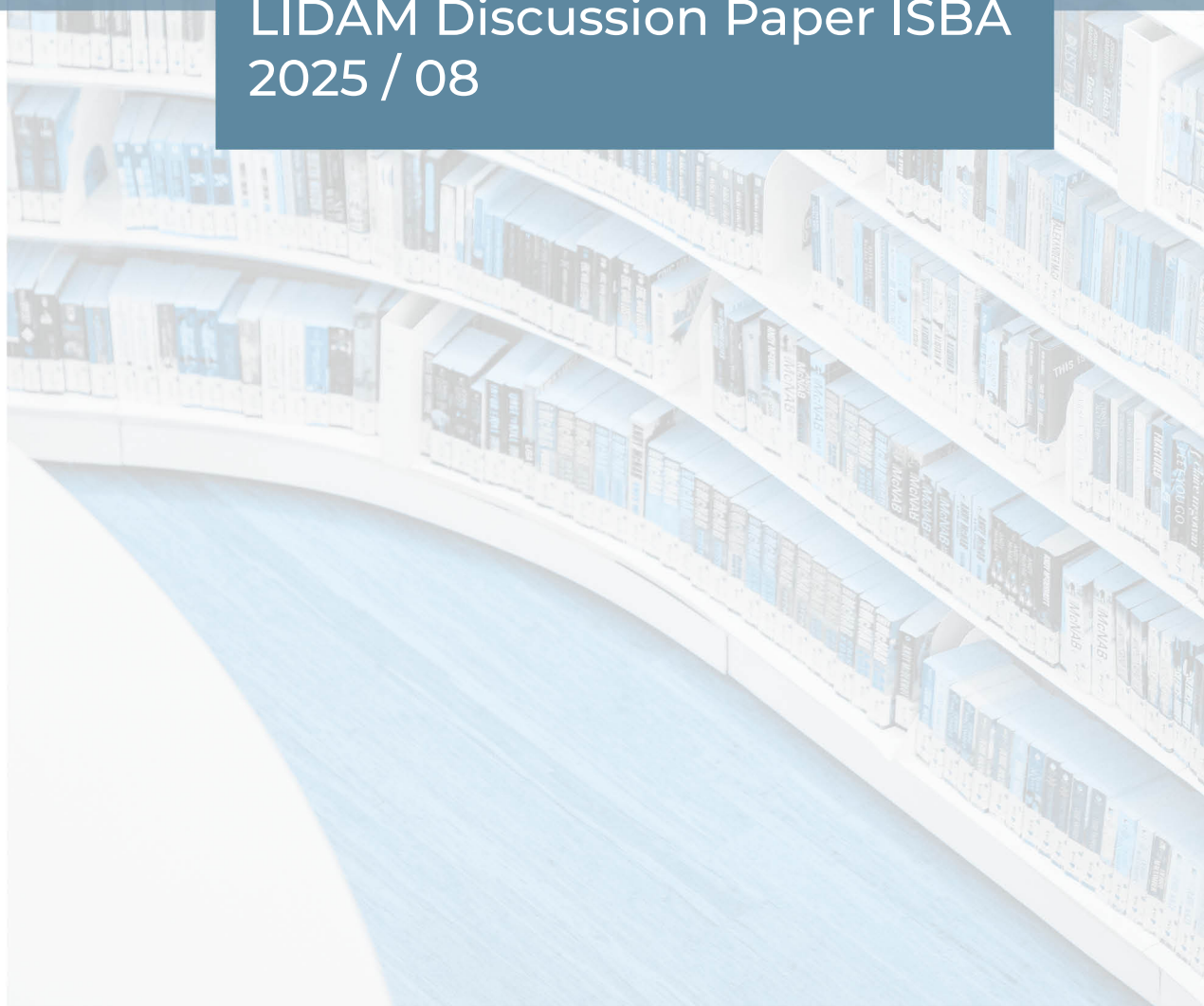
MODELING PRICES FROM SPECULATIVE MARKETS: BURSTING BUBBLES OR DEFLATING BALLOONS?

Christian Hafner, Andrew Harvey, Linqi Wang

LIDAM Discussion Paper ISBA
2025 / 08



UCLouvain

ISBA | LIDAM
Louvain Institute of Data Analysis and Modeling in economics and statistics

ISBA

Voie du Roman Pays 20 - L1.04.01

B-1348 Louvain-la-Neuve

Email : lidam-library@uclouvain.be

https://uclouvain.be/en/research-institutes/lidam/isba/publication

Modeling prices from speculative markets: bursting bubbles or deflating balloons?

Christian M. Hafner*[a], Andrew Harvey†[b], and Linqi Wang‡[b]

[a]Louvain Institute of Data Analysis and modeling in economics and statistics (LIDAM), and ISBA, UCLouvain
[b]Faculty of Economics, University of Cambridge

April 15, 2025

Abstract

Speculative markets may be characterized by sharp falls after a slow build up. Sometimes the converse happens. We suggest a number of mechanisms that are able to produce this kind of behaviour and we demonstrate their plausibility by simulation. The models are then fitted to daily data on Bitcoin. In constructing these models we show that it is essential to take account of volatility and non-normality. We also investigate the possibility of a dynamic tail index. The conclusion, at least for Bitcoin, is that speculative markets are more likely to behave like balloons than bubbles. In other words, there is rapid inflation followed by a slow decline.

Keywords: Bitcoin; explosive models; score-driven models; volatility; tail index.

JEL Classification: C22, C58, G17

*Email: christian.hafner@uclouvain.be.
†Email: ach34@cam.ac.uk.
‡Corresponding author; Email: lw711@cam.ac.uk; Address: Austin Robinson Building, Sidgwick Avenue, Cambridge, CB3 9DD.

Introduction

In a classic bubble, such as the South Sea Bubble or the dot-com bubble, the price of the asset, p_t, shows a steady increase before suddenly collapsing. Modeling such episodes is of limited value because they usually only occur once. The aim here is a more general one, which is to track prices for speculative markets where the fundamentals are weak or non-existent and information is limited. Thus returns may not be martingale differences and as such they may be associated with bubble-like behaviour. The starting point is the Gaussian unobserved components (UC) model:

$$\begin{aligned} y_t &= \mu + \mu_t + \epsilon_t, \quad \epsilon_t \sim NID(0, \sigma_\epsilon^2), \quad t = 1, \ldots, T, \\ \mu_t &= \phi \mu_{t-1} + \eta_t, \quad \eta_t \sim NID(0, \sigma_\eta^2), \qquad |\phi| < 1, \end{aligned} \tag{1}$$

where $y_t = \Delta \ln p_t = \Delta Y_t$. The market is imperfect when $\sigma_\eta^2 > 0$. The above model may be handled by state space methods. However, there are two problems: heteroscedasticity and non-normality. We could set up a parameter-driven model in which ϵ_t is non-normal and subject to a stochastic volatility (SV) mechanism. Dealing with such a model requires a computer intensive procedure as in Durbin and Koopman (2012). The alternative is an observation-driven model. Models in which the dynamics are driven by the conditional score have been developed by Creal et al. (2013), where they were originally called Generalized Autoregressive Score (GAS) models, and Harvey (2013), where they are called Dynamic Conditional Score (DCS) models. We now adopt the term score-driven (SD). SD models for location and scale have been successfully used for a number of years. An early example can be found in Harvey (2013) p 139-41. Here we make some minor extensions, such as using what Blasques et al. (2023) call a quasi score-driven (QSD) model, and allowing the

tail index of a fat-tailed distribution to be time-varying. Models of this kind provide the basis upon which we develop new nonlinear models for location. The aim is to construct dynamic mechanisms that are able to give occasional slow increases in prices followed by sharp falls, that is bubbles,[1] or the opposite, which we call balloons. The sharp increases that characterize a balloon are driven by FOMO - Fear of Missing Out. These models are then fitted to daily observations on Bitcoin returns.

The remainder of the article is organized as follows. Section 1 introduces the two-component time-varying volatility model based on the generalized t distribution, along with an extension incorporating a dynamic tail index. Section 2 presents the score-driven and quasi score-driven location models, while Section 3 develops the time-varying location-scale model and its extension, which better captures locally explosive behaviour in speculative markets. In Section 4, we evaluate the performance of various models in fitting Bitcoin data. Section 5 concludes.

1 Volatility

Before attempting to fit models with bubble and/or balloons we need to establish basic volatility properties under the assumption that y_t is a martingale difference (MD).

1.1 EGARCH

The SD EGARCH model is

$$y_t = \mu + \varepsilon_t \exp(\lambda_{t|t-1}), \quad t = 1, \ldots, T, \tag{2}$$

[1] Blasques et al. (2022a) approach the problem in a different way by introducing thresholds for bubble behaviour.

where the $\varepsilon_t's$ are independently and identically distributed with location zero and unit scale. The stationary first-order dynamic model for $\lambda_{t|t-1}$, the logarithm of the scale, is

$$\lambda_{t+1|t} = \omega^\lambda(1-\phi^\lambda) + \phi^\lambda\lambda_{t|t-1} + \kappa^\lambda u_t, \quad |\phi| < 1, \tag{3}$$

where u_t is the score of the distribution of y_t conditional on past observations, $\lambda_{1|0} = \omega^\lambda$ and ϕ^λ and κ^λ are parameters, with ω^λ denoting the unconditional mean. Letting the conditional distribution be Student t leads to a model known as Beta-t-EGARCH. The EGARCH formulation has the obvious attraction that scale remains positive and that stationarity conditions are straightforward - in (3) we simply require that $\left|\phi^\lambda\right| < 1$. When combined with conditional score dynamics, an asymptotic theory for the ML estimators can be obtained and formulae for moments and autocorrelations may be derived.

Returns may have a different effect on volatility depending on whether they are positive or negative. This effect, sometimes known as leverage, may be modeled by adding another variable to the dynamic equation. Specifically

$$\lambda_{t+1|t} = \omega^\lambda\,(1-\phi^\lambda) + \phi^\lambda\,\lambda_{t|t-1} + \kappa^\lambda\,u_t + \kappa^{\lambda*}u_t^*, \tag{4}$$

where $u_t^* = sgn\,(\mu - y_t)(u_t + 1)$ and $\kappa^{\lambda*}$ is a parameter. A two-component model has often been found to give a significantly improved fit. With leverage it may be written as

$$\begin{aligned} \lambda_{t|t-1} &= \omega^\lambda + \lambda_{1,t|t-1} + \lambda_{2,t|t-1}, \quad t = 1,\ldots,T, \\ \lambda_{i,t+1|t} &= \phi_i^\lambda\,\lambda_{i,t|t-1} + \kappa_i^\lambda\,u_t + \kappa_i^{\lambda*}\,u_t^*, \quad i = 1,2, \end{aligned} \tag{5}$$

where $\phi_1^\lambda > \phi_2^\lambda$ if $\lambda_{1,t|t-1}$ denotes the long-run component. Identifiability requires $\phi_1^\lambda \neq \phi_2^\lambda$, which is implicitly assumed by setting $\phi_1^\lambda > \phi_2^\lambda$, together with the restrictions (i) $\kappa_1^\lambda \neq 0$ or $\kappa_1^{\lambda*} \neq 0$ and (ii) $\kappa_2^\lambda \neq 0$ or $\kappa_2^{\lambda*} \neq 0$.

1.2 Generalized Student t distribution

The use of the generalized Student t distribution gives what Harvey and Lange (2017) call Beta-Gen-t-EGARCH. The probability density function (PDF) for a standardized generalized t variable, symmetric around zero with unit scale, is

$$f(\varepsilon_t) = \frac{\upsilon}{2\eta^{1/\upsilon}} \frac{1}{B(1/\upsilon, \eta/\upsilon)} \left(1 + |\varepsilon_t|^{\upsilon} / \eta\right)^{-(\eta+1)/\upsilon}, \quad -\infty < \varepsilon_t < \infty, \tag{6}$$

where υ and η are positive shape parameters and $B(.,.)$ is the beta function. The variance is given by $\sigma_g^2 = \eta^{2/\upsilon} \frac{\Gamma\left(\frac{3}{\nu}\right)\Gamma\left(\frac{\eta-2}{\nu}\right)}{\Gamma\left(\frac{1}{\nu}\right)\Gamma\left(\frac{\eta}{\nu}\right)}$. Setting $\upsilon = 2$ gives a t distribution with η degrees of freedom, whereas $GED(\upsilon)$ is obtained when $\eta \to \infty$. The Laplace or double exponential distribution is $GED(1)$, while $GED(2)$ is the normal. The parameter η is the tail index, with a range $0 < \eta \leq \infty$, but for both theoretical and practical purposes it is better to work with its inverse, that is $\overline{\eta} = 1/\eta$; in this case $0 \leq \overline{\eta} \leq 1$ for $1 \leq \eta \leq \infty$. The range of υ is $0 < \upsilon < \infty$ but in practice $\upsilon < 1$ is unlikely and it will be assumed that $\upsilon \geq 1$.

In the Beta-Gen-t-EGARCH model, y_t in (2) has a conditional generalized t distribution with mean μ and scale $\exp(\lambda_{t|t-1})$. The conditional score in (3) is

$$u_t = \partial \ln f_t / \partial \lambda_{t|t-1} = (\eta + 1)\, b_t - 1, \tag{7}$$

where $f_t = f(y_t)$ and

$$b_t = \frac{\left(|y_t - \mu| \exp(-\lambda_{t|t-1})\right)^{\upsilon} / \eta}{\left(|y_t - \mu| \exp(-\lambda_{t|t-1})\right)^{\upsilon} / \eta + 1} = \frac{|\varepsilon_t|^{\upsilon}}{|\varepsilon_t|^{\upsilon} + \eta}. \tag{8}$$

is distributed as $beta(1/\upsilon, \eta/\upsilon)$. As $|y_t| \to \infty$, $u_t \to \eta$, so the score is bounded for finite η. Thus it exhibits a soft form of Winsorization. As a result it yields an influence function - or, as it is sometimes called in the volatility literature, news impact curve - that is relatively insensitive to outliers. An exact expression for the information matrix of the

dynamic parameters can be constructed in much the same way as was done for the Beta-t-EGARCH and GB2 location/scale models in Harvey (2013) (sub-sections 4.6 and 5.3). A sufficient condition for invertibility is $\left|\phi^{\lambda} - \kappa^{\lambda}\upsilon(\eta+1)/4\right| < 1$.

Further flexibility in the generalized t can be achieved by introducing skewness and/or creating asymmetry where the shape parameters are allowed to differ on either side of the location parameter. For example, asymmetric Student t distributions were studied by Zhu and Galbraith (2010).

1.3 Time-varying tail models

Financial asset returns often exhibit fatter than normal tails in their distributions which are directly attributable to occasional large absolute price changes. The alternation between periods of normal market activity and episodes of market stress indicates potential time variations in the tail fatness of financial asset returns. In the generalized t distribution, tail fatness is captured by the degrees of freedom η and a dynamic model for time-varying tail parameters may be formulated using a score-driven approach, see Harvey (2022). The dynamics for the log of the inverse degrees of freedom parameter, i.e. $\bar{\eta}^{\dagger} = -\ln\eta = \ln\bar{\eta}$, are modeled by

$$\bar{\eta}^{\dagger}_{t+1|t} = \bar{\eta}^{\dagger}_{t|t-1} + \kappa^{\eta} u_t^{\bar{\eta}^{\dagger}}, \tag{9}$$

where the score is given by

$$u_t^{\bar{\eta}^{\dagger}} = \frac{\partial \ln f_t}{\partial \bar{\eta}^{\dagger}} = \frac{\psi(1/\bar{\eta}\upsilon) - \psi(1/\bar{\eta}\upsilon + 1/\upsilon) + \bar{\eta} - \ln(1-b_t) - (\bar{\eta}+1)b_t}{\upsilon\bar{\eta}}.$$

The process may be initialized with some fixed value $\bar{\eta}^{\dagger}_{1|0}$ which can be treated as an unknown parameter that should be estimated along other parameters.

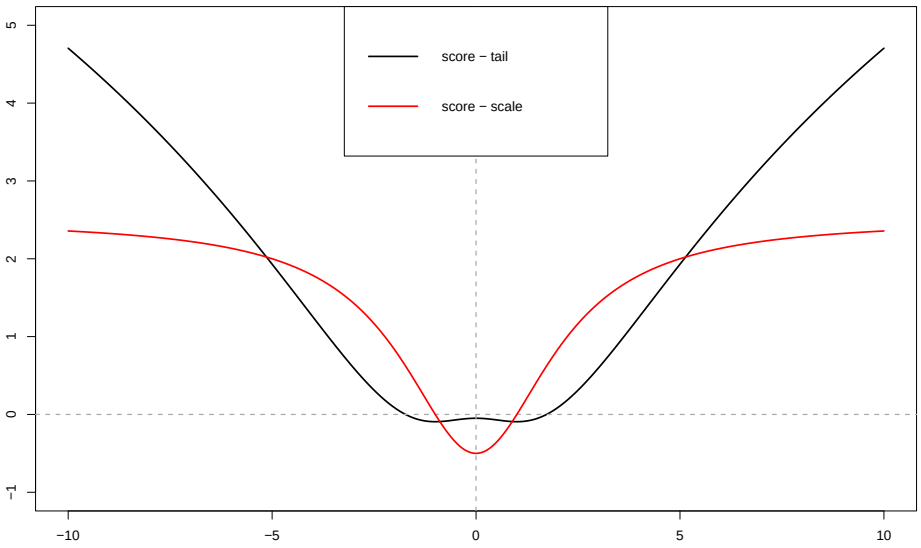


Figure 1: Score for tail v.s. score for scale.

We plot in Figure 1 the score for the tail parameter against the standardized observations y for $\eta = 5$. As y moves toward the tails, $\bar{\eta}^{\dagger}$ increases, so the degrees of freedom fall. This behaviour contrasts with that of the score for the logarithm of scale λ, which is bounded as $y \to \pm\infty$. This is consistent with the intuition that extreme observations should have a larger effect on the tail dynamics than on the scale.

2 Location

The homoscedastic location-t model set up in Harvey and Luati (2014) is

$$y_t = \mu + \mu_{t|t-1} + \varepsilon_t \exp(\lambda), \quad t = 1, \ldots, T, \tag{10}$$

where

$$\mu_{t+1|t} = \phi \mu_{t|t-1} + \kappa u_t^{\mu}. \tag{11}$$

When y_t has a generalized t distribution with conditional mean $\mu + \mu_{t|t-1}$ and scale

$\exp(\lambda)$, the conditional score is

$$u_t^{\mu} = \frac{\partial \ln f}{\partial \mu_{t|t-1}} = \exp(-\lambda)\frac{\eta+1}{\eta}(1-b_t)\,|\varepsilon_t|^{\upsilon-1}\,sgn(\varepsilon_t), \tag{12}$$

where b_t is as in (8) but with $\varepsilon_t = (y_t - \mu - \mu_{t|t-1})\exp(-\lambda)$. The information quantity, as given in Harvey and Lange (2017), is

$$I_{\mu\mu} = \frac{(\overline{\eta}+1)\,\upsilon^3\,\overline{\eta}^{2/\upsilon}}{(\overline{\eta}\,\upsilon+\overline{\eta}+1)\,\exp(2\lambda)}\,\frac{\Gamma\left(2-1/\upsilon\right)\Gamma\left(2/\upsilon+1/\overline{\eta}\upsilon\right)}{\Gamma\left(1/\upsilon\right)\Gamma\left(1/\overline{\eta}\upsilon\right)}, \qquad \upsilon > 1/2.$$

Dividing by $I_{\mu\mu}$ but ignoring constants means that (12) is multiplied by $\exp(2\lambda)$ to give[2]

$$u_t^{\mu} = \exp(\lambda)(1-b_t)\,|\varepsilon_t|^{\upsilon-1}\,sgn(\varepsilon_t) = \exp(\lambda)\frac{|\varepsilon_t|^{\upsilon-1}}{1+\eta^{-1}|\varepsilon|^{\upsilon}}sgn(\varepsilon_t). \tag{13}$$

For a Gaussian distribution, $\eta \to \infty$ and $\upsilon = 2$, so that $u_t^{\mu} = y_t - \mu - \mu_{t|t-1}$. For finite degrees of freedom, the score function is redescending so observations that would be seen as outliers for a Gaussian distribution are far less influential.

2.1 Stationarity and invertibility

The scores u_t^{μ} in (13) are IID and because the score is bounded, the model initialized with non-random starting value for $\hat{\mu}_{t|t-1}$ converges exponentially fast almost surely (e.a.s.) to a unique stationary and ergodic (SE) sequence starting in the infinite past when $|\phi| < 1$, see Blasques et al. (2022b).

For invertibility, we need to express the filter in terms of the observations, whereas stationarity and ergodicity are determined from the model formulated in terms of SE variables. Let $\varphi(\hat{\mu}_{1|0}, y_t, \lambda)$ be the score u_t^{μ} as a function of observables y_t, where $\lambda \in \Lambda$ is the parameter of the innovation density. When $\mu_{t|t-1}$, and hence y_t, is strictly stationary and

[2]The constant factor $\frac{\eta+1}{\eta}$ is dropped as well.

ergodic, and $E\log^+\sup_{\lambda\in\Lambda}\left|\varphi(\hat{\mu}_{1|0},y_1,\lambda)\right|<\infty$, invertibility of the filter is guaranteed, for any $\mu_{1|0}$, by

$$E\log\sup_{\theta\in\Theta}\sup_{\mu^*}\left|\phi+\kappa\frac{\partial\varphi(\mu,y_t,\lambda)}{\partial\mu}\Big|_{\mu=\mu^*}\right|<0 \tag{14}$$

see Proposition 3.2 of Blasques et al. (2022b).

For a standard t-distribution with u_t^μ as in (13) and $\upsilon=2$, we have

$$u_t^\mu=(1-b_t)(y_t-\mu-\mu_{t|t-1}), \tag{15}$$

and so

$$\frac{\partial u_t^\mu}{\partial\mu_{t|t-1}}=(2b_t-1)(1-b_t) \tag{16}$$

It makes no substantive difference if instead we use the raw score or multiply it by $\exp(\lambda)$ rather than $\exp(2\lambda)$ because κ adjusts accordingly. It follows from (16) that the invertibility condition is $\max\{|\phi-\kappa|,\left|\phi+\frac{1}{8}\kappa\right|\}<1$, see Blasques et al. (2022b). The second part is a consequence of the redescending score. The requirement $\phi+\kappa/8<1$ restricts the ability of ϕ to handle persistent processes in which ϕ is close to one.

Similar conditions hold for the generalized t distribution with $\upsilon\geq 2$, see Appendix A. However, in the case of $1\leq\upsilon<2$ - the one of most practical importance - problem arises at $\varepsilon_t=0$. Instead of leading to $|\phi-\kappa|<1$ as for the standard t, $\lim_{\varepsilon\to 0}u'(\varepsilon)=-\infty$, and invertibility cannot be guaranteed. Thus, while a generalized t distribution provides flexibility in capturing fat tails and accommodates a wide range of shapes at the peak of the distribution, the associated SD filter for location may suffer from invertibility issues when $\upsilon<2$. We therefore propose replacing it by a quasi score.

2.2 Quasi score-driven models

The score function with respect to location in the EGB2 distribution exhibits a gentle form of Winsorizing; see Caivano and Harvey (2014). When the scale is parameterized in terms of the standard deviation σ and the distribution is symmetric, we have, after dividing by the information quantity,

$$u_t^{\mu} = \frac{2\xi + 1}{h\xi}\sigma[2b_t - 1], \tag{17}$$

where $h = \sqrt{2\psi'(\xi)}$, $\psi'(\cdot)$ is the trigamma function, and

$$b_t = b_t(\xi) = \frac{\exp\left((y_t - \mu_{t|t-1})h/\sigma\right)}{\exp\left((y_t - \mu_{t|t-1})h/\sigma\right) + 1}.$$

Thus

$$u_t^{\mu} = \frac{2\xi + 1}{h\xi}\sigma\frac{\exp\left((y_t - \mu_{t|t-1})h/\sigma\right) - 1}{\exp\left((y_t - \mu_{t|t-1})h/\sigma\right) + 1} \tag{18}$$

Because $0 \leq b_t \leq 1$, it follows that as $y_t \to \pm\infty$, the score approaches an upper bound of $(2\xi + 1)\sigma/h\xi$.

For a logistic distribution, which has $\xi = 1$, $h = \sqrt{2\psi'(1)} = \pi\sqrt{1/3} \approx 1.814$. As $\xi \to \infty$, $u_t \to y_t - \mu_{t|t-1}$; the mathematics are as in Caivano and Harvey (2014), except that there the forcing variable replacing the score takes a slightly different form.[3] The score-driven model then coincides with the innovations form of the Kalman filter for a Gaussian unobserved components model. At the other extreme, as $\xi \to 0$, the distribution tends to a Laplace and the score is proportional to $sgn(y_t - \mu_{t|t-1})$.

Figure 2 shows the location scores for the three distributions: Gaussian, Student t and logistic. For a logistic distribution, the score is bounded by a constant, while it is redescending towards zero for a Student t distribution for increasing values of the support variable.

[3] To be specific Caivano and Harvey (2014) use $u_t = \sigma h\xi[2b_t - 1]$.

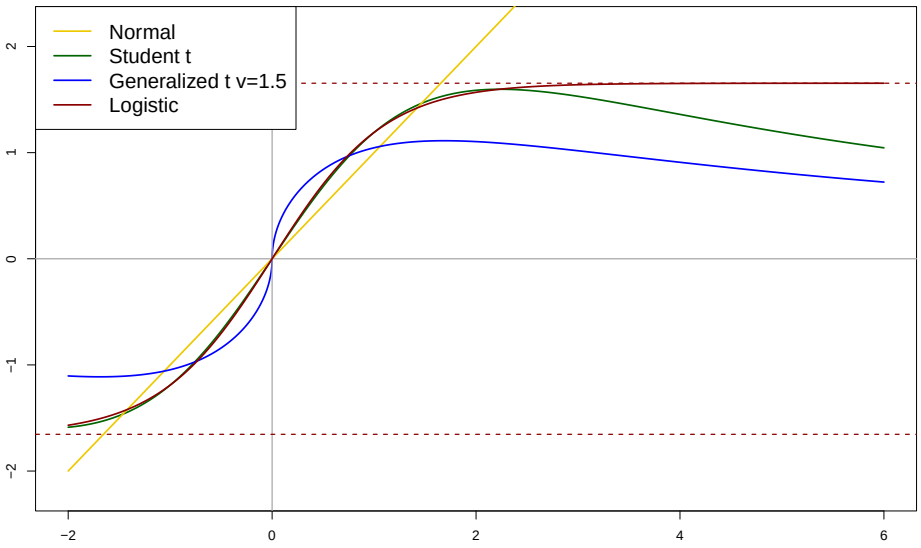


Figure 2: Location scores under different distributions.

The invertibility condition (14) holds if

$$\left|\phi - \kappa \frac{2\xi + 1}{2\xi}\right| < 1; \tag{19}$$

see Appendix A. As with the t-distribution, u_t' does not depend on σ. For a logistic with $\kappa > 0$, $\kappa < 2(1 + \phi)/3$.

The less restrictive invertibility requirements for the EGB2 score make it an attractive alternative to the true score for a generalized t in a QSD model. Under some additional regularity conditions, Theorem 2 of Blasques et al. (2023) shows strong consistency of the MLE towards a pseudo-true parameter, while their Theorem 3 establishes asymptotic normality under possible misspecification.

3 Dynamic location and scale

We now introduce three models that combine score-driven dynamics for both location and scale. In all cases we use a generalized t for the conditional distribution with the corresponding scale score, but a location quasi-score derived from an EGB2 distribution.

3.1 Additive models

A QSD model for dynamic location with an additive EGARCH error may be set up by generalizing (10) to

$$y_t = \mu + \mu_{t|t-1} + \varepsilon_t \exp(\lambda_{t|t-1}), \quad t = 1, \ldots, T, \tag{20}$$

where ε_t has a generalized t distribution. The conditional quasi-score for location, divided by the information quantity, is

$$A1/QSD: \qquad u_t^{\mu} = \frac{2\xi + 1}{h\xi} \sigma_g \exp(\lambda_{t|t-1}) \frac{\exp\left(\varepsilon_t h^{\dagger}\right) - 1}{\exp\left(\varepsilon_t h^{\dagger}\right) + 1}, \tag{21}$$

where $\varepsilon_t = (y_t - \mu - \mu_{t|t-1}) \exp(-\lambda_{t|t-1})$ and $h^{\dagger} = h/\sigma_g$ with σ_g being the standard deviation of generalized t with unit scale. The variable driving the scale is (7), now denoted u_t^{λ}.

The presence of $\exp(\lambda_{t|t-1})$ in the above u_t^{μ} means that it is no longer IID, but if we divide it by the square root of the information quantity, we get

$$A2/QSD: \qquad u_t^{\mu} = \sqrt{2\xi + 1} \frac{\exp\left(\varepsilon_t h^{\dagger}\right) - 1}{\exp\left(\varepsilon_t h^{\dagger}\right) + 1}, \tag{22}$$

which is IID.

3.2 Multiplicative model

Another rationale for dropping $\exp(\lambda_{t|t-1})$ from the forcing variable is to reformulate the model as a multiplicative one,

$$\begin{aligned} y_t &= \mu + \mu_{t|t-1}\exp(\lambda_{t|t-1}) + \varepsilon_t \exp(\lambda_{t|t-1}), \quad t = 1, \ldots, T, \\ &= \mu + \exp(\lambda_{t|t-1})(\mu_{t|t-1} + \varepsilon_t). \end{aligned} \tag{23}$$

The *heteroscedasticity corrected* observations are now modeled by the standard homoscedastic dynamic location model of (10) because

$$(y_t - \mu)\exp(-\lambda_{t|t-1}) = \mu_{t|t-1} + \varepsilon_t, \quad t = 1, \ldots, T. \tag{24}$$

The information quantity for μ in the static model $y_t = (\mu + \varepsilon_t)\exp(\lambda)$ does not depend on λ, suggesting that we use the raw score, that is

$$M/QSD: \qquad u_t^{\mu} = \xi \frac{\exp\left(\varepsilon_t h^{\dagger}\right) - 1}{\exp\left(\varepsilon_t h^{\dagger}\right) + 1}, \tag{25}$$

where $\varepsilon_t = (y_t - \mu)\exp(-\lambda_{t|t-1}) - \mu_{t|t-1}$. Note that here ε_t is not the same as in (21) and (22). Furthermore, the score for $\lambda_{t|t-1}$ now changes to

$$u_t^{\lambda} = (\eta + 1)\, b_t - 1 + \frac{\eta + 1}{\eta}(1 - b_t)\, |\varepsilon_t|^{\upsilon - 1}\, \mu_{t|t-1} sgn(\varepsilon_t). \tag{26}$$

This modification may not be important when $\mu_{t|t-1}$ is dominated by $\lambda_{t|t-1}$.

3.3 Stationarity and invertibility

In model A2/QSD, the location and scale scores u_t^{μ} and u_t^{λ} are IID. Since both scores are bounded, the model initialized with non-random starting values for $\hat{\mu}_{t|t-1}$ and $\hat{\lambda}_{t|t-1}$ converges to a SE series generated by a model initialized in the remote past when $|\phi^{\mu}| <$

1 and $|\phi^{\lambda}| < 1$. For invertibility, while deriving the condition for each component is straightforward, obtaining the condition when both location and scale are time-varying is more challenging given the presence of $\mu_{t|t-1}$ in u_t^{λ} and $\lambda_{t|t-1}$ in u_t^{μ}. In practice, we may resort to simulating the process for different starting values to verify convergence as in, for example, Appendix J of D'Innocenzo et al. (2024).[4] We note that, despite efforts to provide insights into the theoretical properties of the model, establishing the complete estimation theory for score-driven location-scale models remains challenging, as in the ARMA-GARCH case (see, e.g., Francq and Zakoïan (2004); Francq and Zakoïan (2010)).

3.4 Bubbles and balloons

The dynamic location and scale models of the previous section can be extended by introducing further nonlinearities into dynamic location. The aim is to mimic some of the behaviour seen in speculative markets.

When there is a run of positive shocks to returns, $\mu_{t|t-1}$ increases and Y_t increases more rapidly. This feeds speculation leading to a bigger effect on location than can be produced by the models introduced earlier. A dynamic equation for the logarithm of location may be able to capture such behaviour by enhancing the effect of a high level of returns. In a similar way, sharp downward movements may be exaggerated by minus the logarithm of location. Thus $\mu_{t|t-1}$ in (20) or (23) becomes

$$\begin{aligned} \mu_{t|t-1} &= \alpha_1[\exp(\beta_{t|t-1}) - 1] - \alpha_2[\exp(-\beta_{t|t-1}) - 1], \qquad t = 1, \ldots, T, \qquad (27) \\ &= \alpha_2 - \alpha_1 + \alpha_1 \exp(\beta_{t|t-1}) - \alpha_2 \exp(-\beta_{t|t-1}), \qquad \alpha_1, \alpha_2 \geq 0, \end{aligned}$$

[4]For our empirical application, a diagnostic check based on simulations in Appendix C suggests that the filter with multiple time-varying parameters is invertible at the reported estimates.

where

$$\beta_{t+1|t} = \phi^\beta \beta_{t|t-1} + \kappa\, u_t^\beta, \quad \kappa \geq 0,\ 0 \leq \phi^\beta \leq 1. \tag{28}$$

The starting value is $\beta_{1|0} = 0$. (Note that $\kappa > 0$ unless ϕ^β is fixed.) When $\beta_{t|t-1} = 0$, $\exp(\beta_{t|t-1}) = \exp(-\beta_{t|t-1}) = 1$ which is the reason for subtracting unity from the exponents in (27).[5]

When $\alpha_2 = 0$, the response induced by (27) is as shown by the blue dashed line in Figure 3. Large positive movements in $\beta_{t|t-1}$ are amplified, leading to a faster rise in Y_t. There is no corresponding rapid fall: hence we have a balloon. On the other hand, when $\alpha_1 = 0$, the opposite happens. This is a bubble. When $\alpha_1 = \alpha_2 = \alpha$, the model can be written as

$$y_t = \mu + 2\alpha \sinh(\beta_{t|t-1}) + \varepsilon_t \exp\left(\lambda_{t|t-1}\right)$$

where $\sinh(x) = (e^x - e^{-x})/2$ and we denote this model as bubble/balloon [sinh]. This is the green solid line in Figure 3. In this case, large movements of $\beta_{t|t-1}$ in either direction are amplified, leading to pronounced rises and falls in Y_t. When $\alpha_1 = \alpha_2 = 0$ the mean in (27) reduces to zero and y_t is a MD. (Testing the hypothesis $\alpha_1 = \alpha_2 = 0$ is however non-standard because the parameters lie on a boundary.)

The score with respect to $\beta_{t|t-1}$ in (27) is

$$\begin{aligned} \frac{\partial \ln f_t}{\partial \beta_{t|t-1}} &= \frac{\partial \ln f_t}{\partial \mu_{t|t-1}} \frac{\partial \mu_{t|t-1}}{\partial \beta_{t|t-1}} = \frac{\partial \ln f_t}{\partial \mu_{t|t-1}} \left\{ \alpha_1 \exp\left(\beta_{t|t-1}\right) + \alpha_2 \exp\left(-\beta_{t|t-1}\right) \right\} \\ &= u_t^\mu \exp(-\lambda_{t|t-1}) \left\{ \alpha_1 \exp\left(\beta_{t|t-1}\right) + \alpha_2 \exp\left(-\beta_{t|t-1}\right) \right\}, \end{aligned}$$

where u_t^μ is the score in (22) for A2. Dividing by the square root of the (conditional) information quantity yields $u_t^\beta = u_t^\mu$, which is IID.

[5] It would in principle be better to subtract the expectation of the exponential terms but this is unduly complicated.

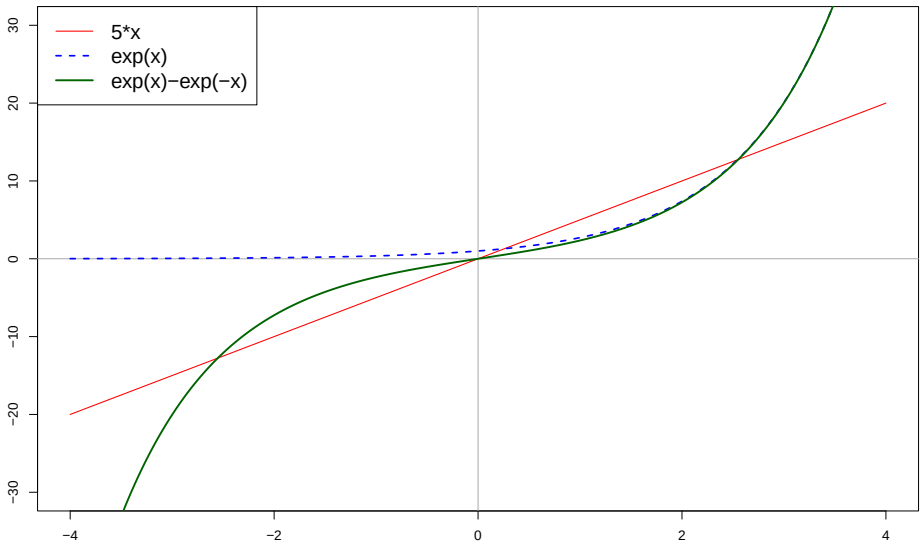


Figure 3: The location parameter as a function of $\beta_{t|t-1}$ in the balloon model (dashed blue curve) and bubble/balloon model (thick green curve). For reference we show a linear function (thin red) as in the additive or multiplicative linear location models.

For the homoscedastic model, the sufficient condition for invertibility holds for the quasi score-driven model with EGB2 score dynamics, provided that

$$\left|\phi - \kappa\overline{\alpha}(\exp(\bar{\beta}) + \exp(-\bar{\beta}))\frac{2\xi+1}{2\xi}\right| < 1, \tag{29}$$

where $\overline{\alpha} = \max(\alpha_1, \alpha_2)$, $\bar{\beta} = \frac{\kappa}{1-\phi}\bar{u}$, where $\bar{u}$ is max on $\left|u_t^{\mu}\right|$. See Appendix B for a proof.

Note that for the special case of a logistic distribution, i.e. $\xi = \varsigma = 1$, (29) simplifies to $\left|\phi - \frac{3}{2}\kappa\bar{\alpha}(\exp(\bar{\beta}) + \exp(-\bar{\beta}))\right| < 1$.

3.5 Simulation examples

We simulate the bubble and balloon models to illustrate how they can generate some of the behaviour observed in speculative markets. We focus on two special cases: i) the

exponential transformation with $\alpha_1 = 0.06$ and $\alpha_2 = 0$; and ii) the sinh transformation with $\alpha_1 = \alpha_2 = 0.06$. We use the parameter estimates obtained from our empirical application for the volatility and location processes. The simulated series are plotted in Figures 4 and 5, with the upper panel in each figure being the log price series and the lower panel being the conditional mean. From Figure 4, we observe in the lower panel that the exponential transformation amplifies the large positive values of the β_t process and generates explosive behaviour during periods of rapid price increases. However, downward adjustments tend to proceed more slowly. From Figure 5, we can see that the sinh transformation amplifies both large positive and negative values of the β_t process, leading to rapid upward and downward movements in prices.

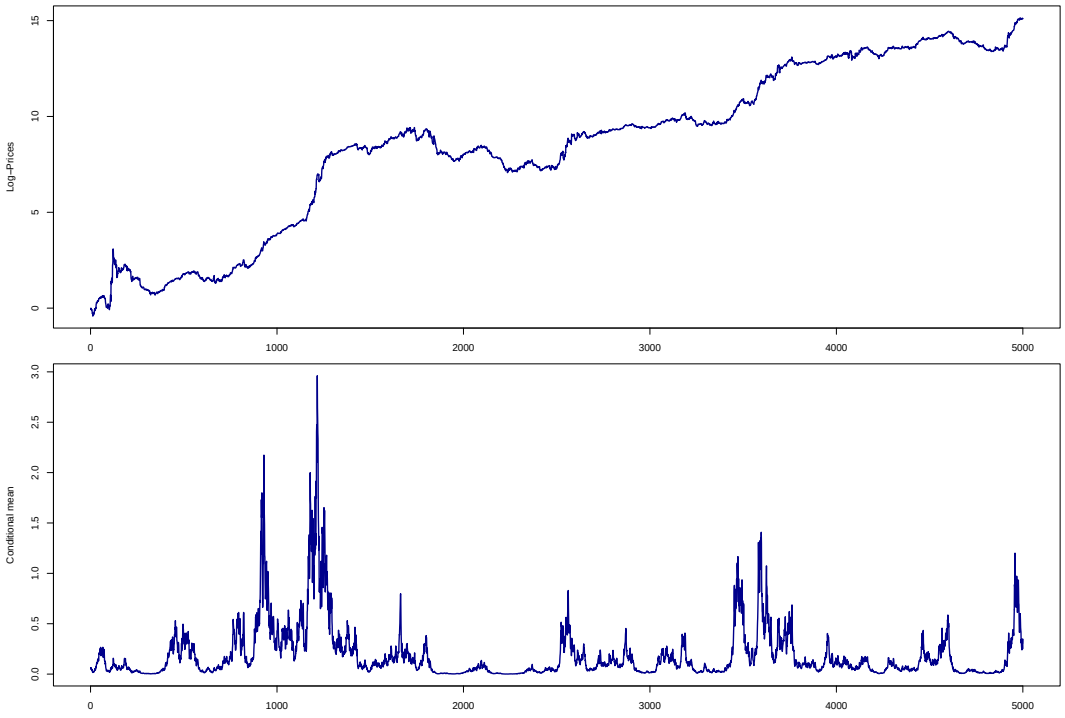


Figure 4: Simulation example of QSD-balloon model $\alpha_1 = 0.06, \alpha_2 = 0$.

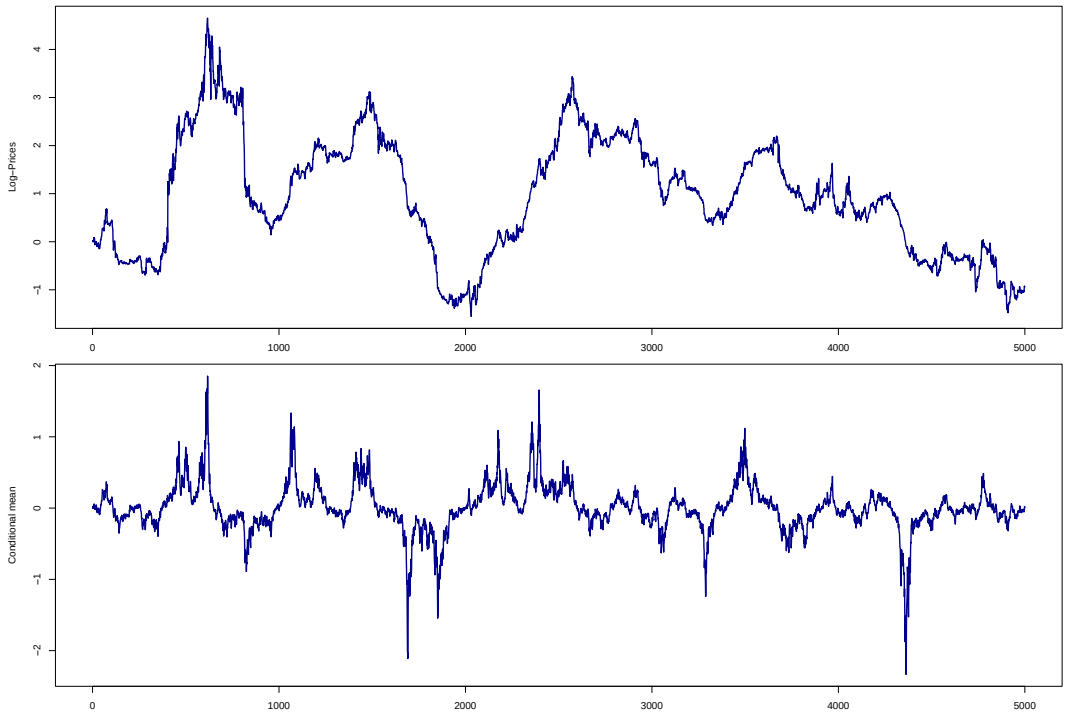


Figure 5: Simulation example of QSD-bubble/balloon [sinh] model $\alpha_1 = \alpha_2 = 0.06$.

4 Application to Bitcoin

In our application, we consider the daily return series of Bitcoin, a well-documented speculative asset that has experienced multiple episodes of extreme volatility and bubbles (Hafner (2020)). The sample period spans from 19 July 2010 to 12 November 2024, with a total of 5,231 daily observations.[6] Figure 6 presents the price, log price, and return series. We observe several episodes with significant price surges, followed by abrupt market corrections with drawdowns of 80-90% taking place over a short horizon. The high level of volatility in Bitcoin returns in the early part of the sample (2010-2015) reflects the low liquidity of the market in its early phase, with most of Bitcoin trading against US dollar taking place

[6]The data is retrieved from Coin Metrics.

through the exchange platform Mt. Gox (Scaillet et al. (2020)).[7] We see an exponential growth phase from 2015 to end-2017, supported by increasing media coverage and the introduction of derivative contracts on Bitcoin by traditional exchanges (Hafner (2020)). The high failure rate of newly launched cryptocurrencies, combined with several countries taking an aggressive stance against them, led to a reversal in market sentiment and a sharp correction over 2018. Finally, the growth of Bitcoin in the more recent part of the sample is supported by the rise of decentralized finance (DeFi) platforms and stablecoins which allow traders to build leveraged positions in other cryptocurrencies (Gorton et al. (2022)). However, the increasing use of leverage in crypto assets also leads to sharper market corrections during periods of heightened volatility as margin calls fuel deleveraging pressures. An illustration of this mechanism is the market meltdown of 2022.

[7]Indeed, several of the market corrections observed during this period can be traced back to cybersecurity incidents which eventually led the exchange to suspend its operations and declare bankruptcy in early 2014.

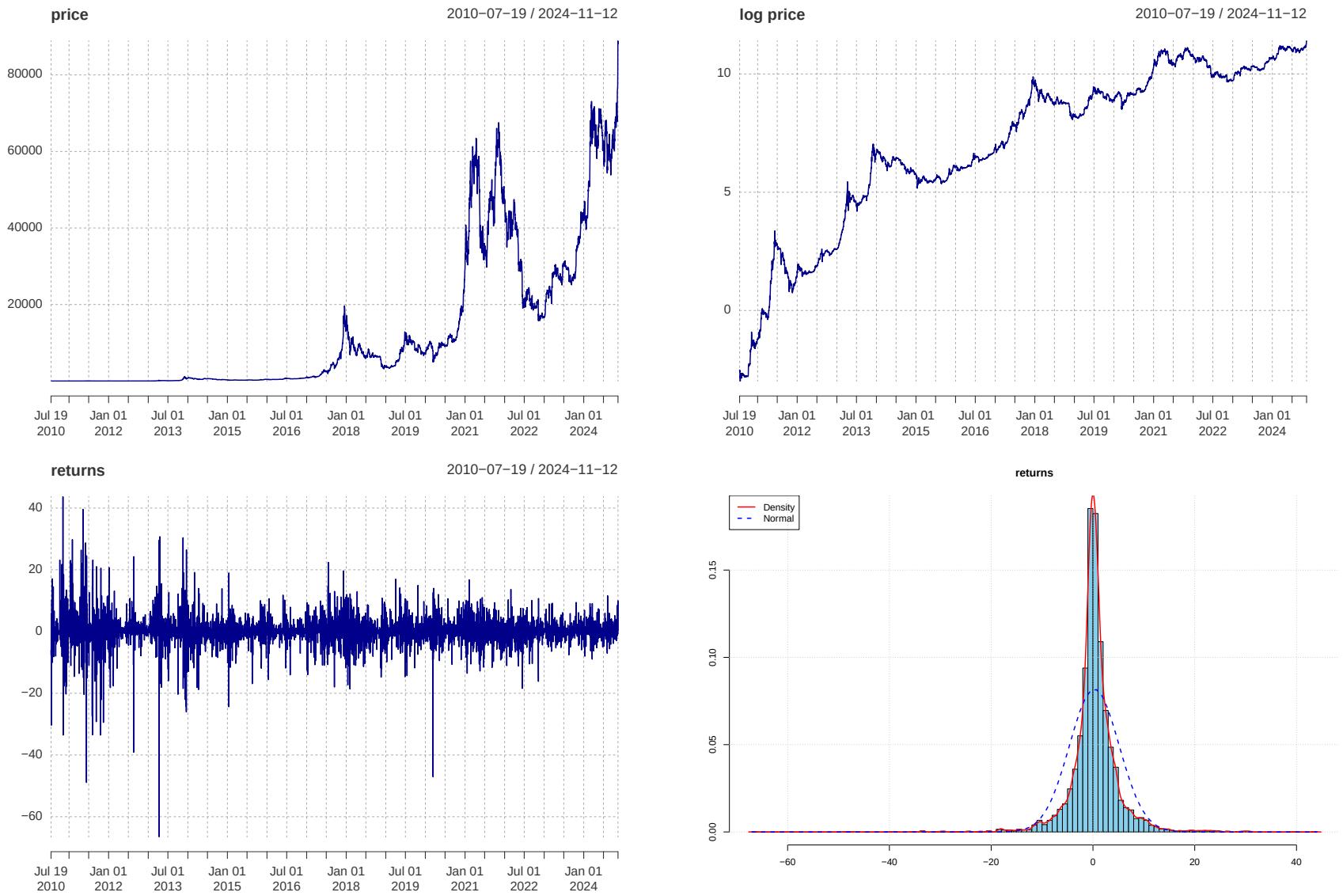


Figure 6: Daily Bitcoin price (USD), log price, and returns.

Descriptive statistics of the Bitcoin return series are provided in Table 1. The series exhibits a positive mean, high volatility, strong negative skewness and high kurtosis over the sample period. The extreme behaviour of cryptocurrencies compared to more traditional assets such as equity has been highlighted previously, see e.g. Catania and Grassi (2022). These observations warrant the use of filtering methods robust to extreme observations, combined with a fat-tailed distribution, to better capture the key features of the data. The sharp peak in the histogram of returns supports the use of generalized t distribution.

Table 1: Descriptive statistics for the daily log-return series of Bitcoin.

	T	Min	Median	Max	Mean	StdDev	Skewness	Kurtosis
Bitcoin	5231	-66.495	0.162	43.666	0.265	4.876	-0.750	20.177

4.1 Long-run and short-run volatilities

Our baseline model is the score-driven two-component volatility model with a generalized t conditional distribution (EGARCH-2C). The specification allows for different asymmetric responses (leverage effects) in the short and long run. This choice is motivated by mounting evidence in the literature that these models are a robust benchmark for volatility modeling.

Table 2: Two-component Beta-Gen-t-EGARCH model estimates.

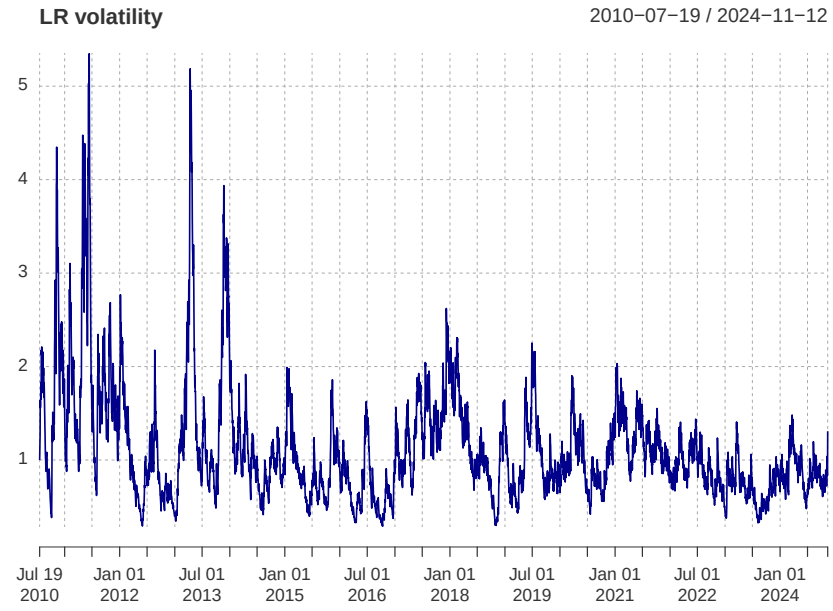
		$1/\eta$	υ	ω^λ	κ_1^λ	ϕ_1^λ	$\kappa_1^{\lambda*}$	κ_2^λ	ϕ_2^λ	$\kappa_2^{\lambda*}$	μ	$\ln L$	AIC	BIC
EGARCH-2C	est.	0.174	1.193	0.825	0.089	0.981	-0.015	0.076	0.638	0.040	0.108	-13645.14	27310.29	27375.91
	s.e.	0.029	0.066	0.099	0.012	0.004	0.005	0.019	0.103	0.010	0.029			

Note: Number of parameters $k = 10$.

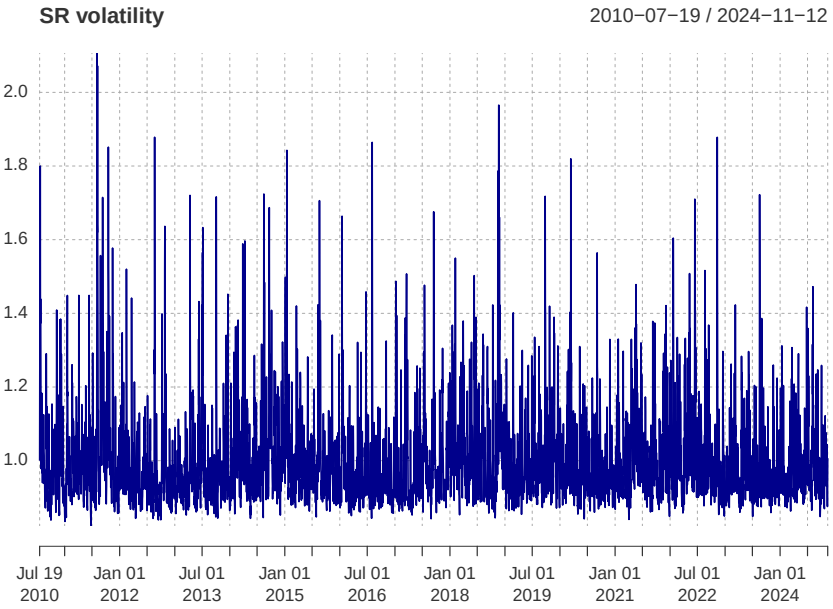
We present the parameter estimates for the EGARCH-2C model in Table 2, together with their associated standard errors. In addition, we report the log-likelihood, Akaike Information Criterion (AIC), and Bayesian Information Criterion (BIC). The autoregressive coefficient ϕ_1^{λ} is close to one indicating high persistence in the long-run volatility process. It is significantly different from the coefficient estimate for the short-run volatility process, which is 0.64. The estimated $\kappa_1^{\lambda *}$ parameter is negative and significant, though its magnitude is relatively small. In contrast, for the short-term news impact curve, $\kappa_2^{\lambda *}$ is positive and twice as large. This suggests that negative returns increase short-run volatility more than positive returns, in line with evidence for other asset classes; see, for example Harvey and Lange (2018); Harvey and Palumbo (2023).[8] Moreover, the estimated $\bar{\eta}$ parameter is 0.17, corresponding to a tail index between 5 and 6. Lastly, the estimate of the additional parameter υ for the generalized t distribution is 1.19, which is significantly different from 2, its value in a standard Student t distribution.

Figure 7 plots the evolution of the filtered long and short-run volatility components. For the long-run component, we observe a decrease in volatility over the studied period as low market liquidity is associated with higher levels of volatility in the early part of the sample. In the second part of the sample, we observe spikes in the long-run volatility component at each turning point between the end of a rally period and the start of a market correction. For the short-run component, we observe that volatility tends to be higher during periods of pronounced market drawdowns, in line with the leverage effect

[8]The different responses of long-run and short-run volatility to news could make the leverage effect appear insignificant or negative in a one-component model, further supporting the choice of a two-component volatility model. This observation also provides a potential reconciliation with evidence in previous literature of an "inverted" leverage effect for the volatility of Bitcoin returns (Ardia et al. (2019)).



(a) Long-run volatility.



(b) Short-run volatility.

Figure 7: Time-varying volatility

and our parameter estimate for $\kappa_2^{\lambda*}$ in Table 2.[9]

[9]See e.g. the periods from December 2013 to December 2015 and from December 2017 to December 2018 where the price of Bitcoin dropped by more than 80%.

4.2 SD versus QSD models

We next turn to the additive models allowing for time-varying locations with a two-component EGARCH volatility model, namely the A1 and A2 models introduced in Section 3.1. In addition, we explore the potential benefits of the QSD location filter derived from the logistic distribution.

Table 3 presents the parameter estimates and goodness of fit statistics for the A1 and A2 models with SD and QSD dynamics. First, we observe that the volatility parameter estimates remain quite stable across the different specifications of the location model. The dynamic location process exhibits high persistence, as evidenced by the estimated autoregressive parameters $\hat{\phi}$ being very close to one in all cases. In terms of model fit, we observe that both the SD and QSD A2 models provide a better fit than the A1 models, so it seems that using an IID score provides better empirical performance. Lastly, the A1/QSD and A2/QSD models deliver improved log-likelihood values compared to their SD counterparts. The model performance also improves in terms of lower AIC and BIC statistics as no additional parameters are introduced by the logistic score function.

4.3 Locally explosive models

Given the superior performance of the QSD models over their SD counterparts, we focus on QSD models that can mimic the locally explosive behaviour seen in the Bitcoin series. We consider the multiplicative location-scale model (M/QSD) introduced in Section 3.2, and extensions of the additive model featuring a nonlinear transformation as specified in (27) of Section 3.4. Unrestricted estimation gave an estimate of α_2 equal to zero, so we focus on two special cases: $\alpha_2 = 0$, corresponding to an exponential transformation (QSD-balloon),

Table 3: Estimation results using SD and QSD additive location-scale models.

		$1/\eta$	υ	ω^λ	κ_1^λ	ϕ_1^λ	$\kappa_1^{\lambda*}$	κ_2^λ	ϕ_2^λ	$\kappa_2^{\lambda*}$	μ	κ	ϕ	$\ln L$	AIC	BIC
A1/SD	est.	0.182	1.218	0.833	0.084	0.983	-0.017	0.078	0.677	0.038	0.124	0.010	0.978	-13639.63	27303.25	27382.00
	s.e.	0.017	0.033	0.078	0.011	0.004	0.004	0.015	0.073	0.008	0.011	0.001	0.002			
A2/SD	est.	0.182	1.213	0.832	0.087	0.982	-0.015	0.079	0.636	0.039	0.117	0.017	0.993	-13635.93	27295.87	27374.62
	s.e.	0.023	0.053	0.064	0.011	0.004	0.005	0.018	0.085	0.011	0.007	0.001	0.001			
A1/QSD	est.	0.174	1.195	0.824	0.083	0.983	-0.017	0.081	0.678	0.040	0.102	0.010	0.978	-13637.86	27299.71	27378.46
	s.e.	0.003	0.012	0.006	0.002	0.002	0.003	0.004	0.002	0.005	0.003	0.004	0.002			
A2/QSD	est.	0.174	1.190	0.826	0.086	0.983	-0.016	0.079	0.663	0.040	0.099	0.031	0.985	-13632.72	27289.44	27368.19
	s.e.	0.010	0.009	0.017	0.012	0.004	0.002	0.019	0.012	0.005	0.006	0.006	0.004			

Note: Number of parameters $k = 12$ for all models.

and $\alpha_1 = \alpha_2$, corresponding to a sinh transformation (QSD-bubble/balloon).

Table 4 presents the parameter estimates along with the log-likelihood, AIC, and BIC. First, we note that all models outperform the additive models (A1 and A2) presented in Table 3, as evidenced by higher log-likelihood and lower AIC and BIC. Furthermore, the QSD-balloon provides the best fit as the exponential transformation facilitates timely and substantial updates during price surges which have occurred multiple times for Bitcoin. Lastly, the M/QSD model achieves the second-best performance. By incorporating the volatility process into the specification of the time-varying location parameter, it allows the conditional mean to update more during periods of high volatility.

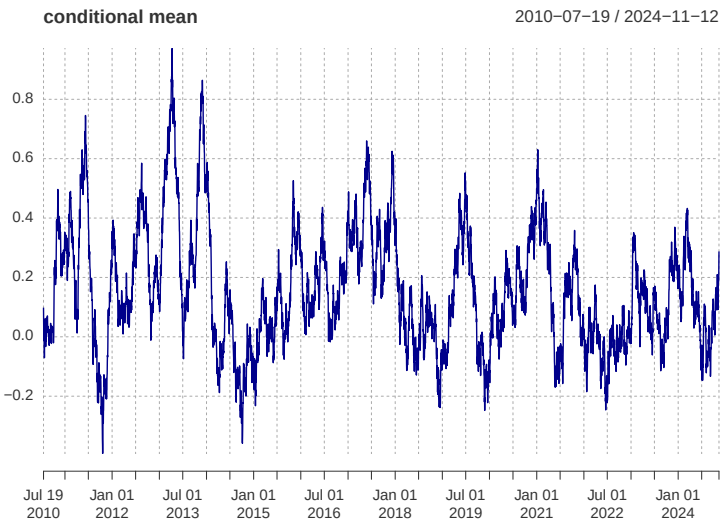
We plot in Figure 8 the estimated time-varying conditional mean for the different models. We first observe that the handful of large outliers (see Figure 6) do not materially impact the filtered conditional mean estimates because the score-driven filter downweights their influence when updating the time-varying location parameter. This highlights the

Table 4: Estimation results using QSD location/bubble/balloon-volatility models.

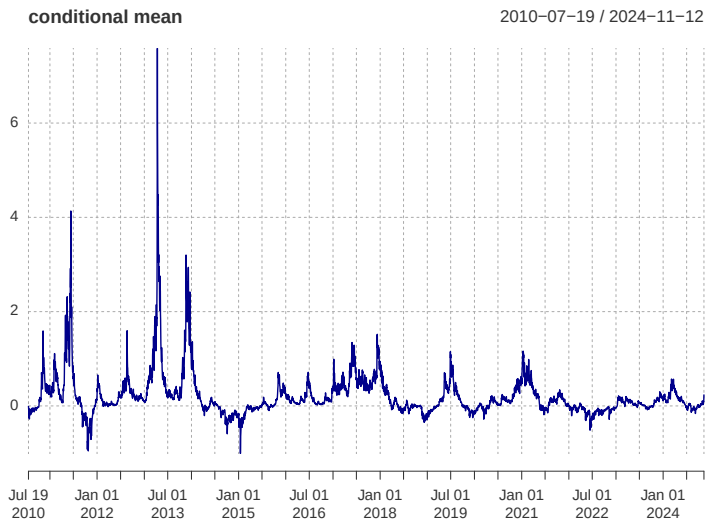
		$1/\eta$	υ	ω^λ	κ_1^λ	ϕ_1^λ	$\kappa_1^{\lambda*}$	κ_2^λ	ϕ_2^λ	$\kappa_2^{\lambda*}$	α	κ	ϕ	$\ln L$	AIC	BIC
M/QSD	est.	0.171	1.177	0.818	0.088	0.983	-0.017	0.081	0.657	0.042		0.016	0.996	-13629.90	27281.81	27353.99
	s.e.	0.008	0.006	0.006	0.005	0.002	0.003	0.013	0.013	0.002		0.002	0.001			
QSD-balloon	est.	0.174	1.189	0.839	0.086	0.984	-0.015	0.080	0.674	0.038	0.059	0.201	0.990	-13624.99	27273.98	27352.72
	s.e.	0.005	0.005	0.005	0.004	0.003	0.002	0.011	0.005	0.009	0.003	0.006	0.002			
QSD-bubble/balloon [sinh]	est.	0.172	1.181	0.828	0.087	0.984	-0.013	0.081	0.665	0.039	0.009	0.195	0.996	-13627.43	27278.86	27357.61
	s.e.	0.020	0.040	0.047	0.011	0.004	0.005	0.014	0.034	0.009	0.010	0.016	0.001			

Note: Number of parameters $k = 11$ for the M/QSD model and $k = 12$ for the QSD-balloon and QSD-bubble/balloon [sinh] models.

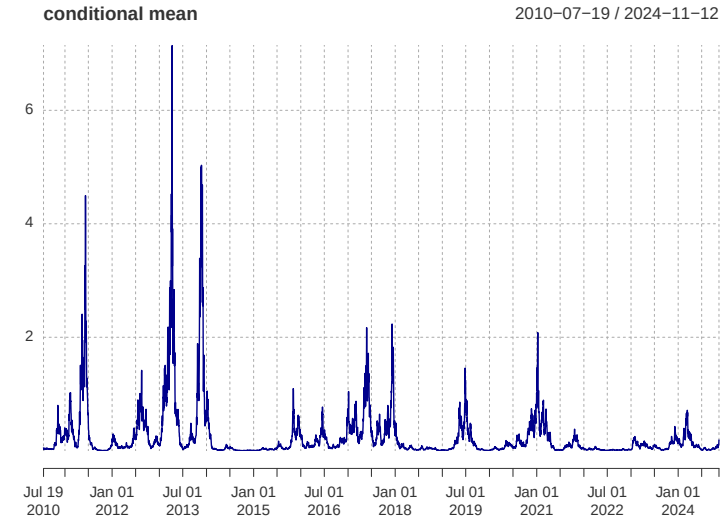
attractiveness of score-driven models for the modeling of highly volatile series such as cryptocurrencies which experience frequent episodes of market rallies and crashes. Second, the conditional mean levels of the different models track well the significant spikes occurring during explosive episodes and the sharp drops during market corrections. The QSD-balloon model appears to capture the frequency and sharper spikes more effectively thanks to the exponential transformation introduced in (27). The M/QSD model, which features an interaction term between the time-varying location and scale parameters, see (23), tracks the spikes in the conditional mean more closely in the earlier period when Bitcoin experienced higher volatility. However, the model's ability to capture the higher conditional mean during speculative episodes is weakened in the second part of the sample as the improvement in market liquidity results in lower volatility levels overall. Accounting for these different aspects, the QSD-balloon model appears to be better suited to modeling the conditional mean dynamics of Bitcoin returns because it can capture the intensity of speculative surges and market corrections across episodes of both high and low market volatility.



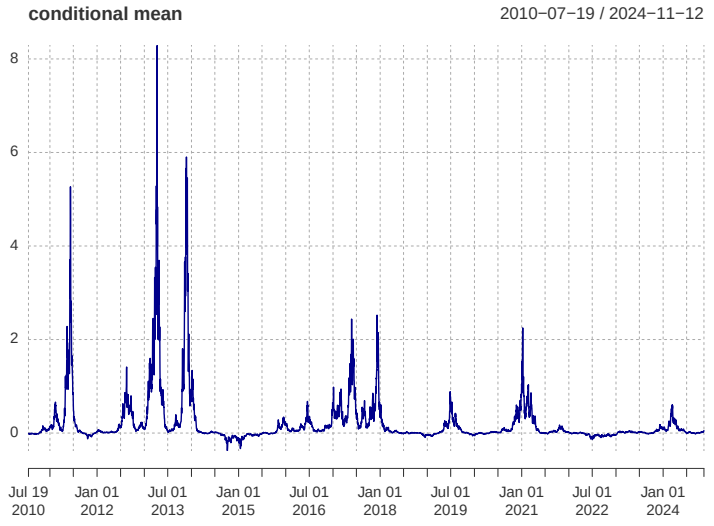
(a) A2/QSD



(b) M/QSD



(c) QSD-balloon



(d) QSD-bubble/balloon [sinh]

Figure 8: Time-varying conditional mean.

4.4 Time-varying tails

Figure 6 suggests that the occurrence of extreme observations in Bitcoin returns varies over time, with larger and more frequent extremes being observed in the earlier part of the sample. To capture these time variations in tail behaviour, we extend the QSD-balloon models to incorporate time-varying tail fatness in the conditional distribution (see Section 1.3). To isolate the tail behaviour from the volatility dynamics, we fix the parameters for the time-varying location and scale processes to their estimates from the QSD-balloon model and estimate the additional parameters for the tail dynamics.

We plot in Figure 9 the filtered tail index, which varies mostly between approximately four and seven. In line with previous evidence, we see that the tail fatness is particularly high at the beginning of the sample when market liquidity was lower. Prolonged explosive episodes are associated with an increase in the tail index, that is, a decrease in tail fatness. This upward trend over the studied sample is interspersed with periodic drops in the tail index during sudden market crashes.

4.5 Out-of-sample forecasting performance

In this section, we evaluate the out-of-sample performance of the various models during the well-documented speculative bubble period of 2017-2018. Our analysis focuses on one-step-ahead density forecasts of y_{T+1}. We use the return series up to February 2017 as the in-sample data to estimate various models, and generate rolling one-step-ahead forecasts of the location and volatility for the out-of-sample forecast period ending in March 2018. Under the assumption of a conditionally generalized t distribution, we construct density forecasts using the one-step-ahead predictions of the time-varying parameters.



Figure 9: Time-varying tail index series.

To assess the out-of-sample model performance, we start with a widely used scoring rule—the average negative log score (NLS)—which is defined as

$$\overline{\text{NLS}} = -\log p\left(y_{T+1}; \widehat{\boldsymbol{\theta}}_{T+1|T}\right) \tag{30}$$

where $\theta_{T+1|T} = (\mu_{T+1|T}, \lambda_{T+1|T})$. A lower value of NLS indicates better out-of-sample predictive accuracy. We use the two-component EGARCH model (with constant location) as our baseline and plot in Figure 10 the one-step ahead forecast of the conditional mean alongside the cumulative sum of the differences in log scores between models incorporating time-varying location and the EGARCH model for the period of 2017–2018. The inclusion of time-varying location parameters results in notable improvements in the forecasting performance. This is reflected in the positive and generally upward-sloping cumulative log score differences of the A2/QSD, M/QSD, and QSD-balloon models relative to the EGARCH model—indicating higher predictive accuracy.[10] Furthermore, the exponential

[10]The performance of the QSD-bubble/balloon [sinh] is very close, but inferior, to the one of the QSD-

transformation employed in the QSD-balloon model captures higher-frequency fluctuations and sharper spikes more effectively during the out-of-sample period, highlighting its consistent performance in modeling explosive dynamics for the Bitcoin series both in and out-of-sample.

We also consider the average weighted continuous ranked probability score (wCRPS) criterion as an alternative scoring rule, which is defined as

$$\overline{\text{wCRPS}} = \int_{-\infty}^{\infty} w(z) \left(F\left(z; \widehat{\boldsymbol{\theta}}_{T+1|T}\right) - \mathbb{I}_{\{y_{T+1}<z\}} \right)^2 \, \mathrm{d}z \ , \tag{31}$$

where $w(z)$ is a weight function that emphasizes specific regions of the predictive distribution, such as the tails or the center, and $F\left(z; \hat{\boldsymbol{\theta}}_{T+1|T}\right)$ is the predictive cumulative density function evaluated at z. We consider different weighting schemes emphasizing either the whole distribution (uniform), the center or the tails—see Table 5.

Table 5: Weighting functions.

Emphasis	weight function
Uniform	$w(z) = 1$
Center	$w(z) = \phi_{a,b}(z)$
Tails	$w(z) = 1 - \phi_{a,b}(z)/\phi(a)$
Right tail	$w(z) = \Phi_{a,b}(z)$
Left tail	$w(z) = 1 - \Phi_{a,b}(z)$

Note: $\phi(z)$ and $\Phi(z)$ denote respectively the pdf and cdf of a $N(0,1)$ distribution.

Table 6 provides the test statistics for Diebold and Mariano (2002) tests of equal predictive accuracy in terms of the wCRPS criterion between our proposed models and the balloon model.



Figure 10: One-step ahead forecast of the conditional mean and the cumulative sum of the differences in log scores between models with and without time-varying location.

EGARCH benchmark model. The columns correspond to the different weighting schemes focusing on specific parts of the distribution. The negative values indicate that the models including a time-varying location generate more accurate predictions of the one-step ahead conditional distribution than the EGARCH model. In addition, the improvement in forecasting accuracy appears to be more pronounced in the center of the distribution and in the right tail as the test statistics are significantly different from zero in all cases.

Table 6: Test statistics for the Diebold and Mariano (2002) test of equal performance between the series of weighted continuous ranked probability scores for various models and the EGARCH model.

	uniform	center	tails	right tail	left tail
A2/QSD	-2.37^{***}	-2.63^{***}	-1.51^{*}	-2.45^{***}	-2.17^{**}
M/QSD	-1.08	-1.31^{*}	-0.33	-1.34^{*}	-0.74
QSD-balloon	-1.07	-1.41^{*}	-0.25	-1.45^{*}	-0.65
QSD-bubble/balloon [sinh]	-1.10	-1.43^{*}	-0.38	-1.46^{*}	-0.70

Note: Rejection of the null hypothesis of equal predictive ability at the 1%, 5% and 10% confidence levels are represented respectively by ***, ** and *.

5 Conclusion

In this paper, we propose a set of quasi score-driven models for time-varying location that are designed to capture the locally explosive behaviour often encountered in speculative markets. Our empirical study on Bitcoin data yields five key findings. First, adopting a two-component score-driven EGARCH volatility model—accounting for both short- and

long-term volatility dynamics—is essential for accurately modeling the time series of Bitcoin returns over a long sample period. Second, the QSD models for dynamic location provide less restrictive invertibility conditions and a better overall fit compared to their score-driven counterparts. Third, employing the exponential transformation in the location model allows for more accurate tracking of the price dynamics during periods of rapid increases, demonstrating the model's ability to effectively capture sudden surges in speculative markets. These tend to be followed by a gradual decline, more reminiscent of a deflating balloon rather than a bursting bubble. Fourth, incorporating time-varying tail dynamics further enhances the capacity of the models to accommodate the large and more frequent extreme observations in the Bitcoin return series during the earlier part of the sample. Finally, we confirm in an out-of-sample forecasting exercise that the models with time-varying location dynamics offer attractive improvements in density forecast accuracy compared to the EGARCH benchmark.

Acknowledgements

We are grateful for helpful comments and discussions with seminar participants at the Center for Research in Economics and Statistics (CREST), at the 16th CMStatistics conference December 2023 and at the Tinbergen Institute Time Series Econometrics Theory Workshop. The first author gratefully acknowledges financial support from the Belgian Federal Science Policy (contract ARC 18/23-089). The second author thanks the University of Cambridge Keynes Fund for support on the project Persistence and Forecasting in Climate and Environmental Econometrics, and the third author acknowledges financial support from the Janeway Institute.

References

Ardia, D., Bluteau, K. and Rüede, M. (2019): "Regime changes in Bitcoin GARCH volatility dynamics". *Finance Research Letters* 29, pp. 266–271.

Blasques, F., Francq, C. and Laurent, S. (2023): "Quasi score-driven models". *Journal of Econometrics* 234.1, pp. 251–275.

Blasques, F., Koopman, S. J. and Nientker, M. (2022a): "A time-varying parameter model for local explosions". *Journal of Econometrics* 227.1, pp. 65–84.

Blasques, F., van Brummelen, J., Koopman, S. J. and Lucas, A. (2022b): "Maximum likelihood estimation for score-driven models". *Journal of Econometrics* 227.2, pp. 325–346.

Caivano, M. and Harvey, A. (2014): "Time-series models with an EGB2 conditional distribution". *Journal of Time Series Analysis* 35.6, pp. 558–571.

Catania, L. and Grassi, S. (2022): "Forecasting cryptocurrency volatility". *International Journal of Forecasting* 38.3, pp. 878–894.

Creal, D., Koopman, S. J. and Lucas, A. (2013): "Generalized autoregressive score models with applications". *Journal of Applied Econometrics* 28.5, pp. 777–795.

Diebold, F. X. and Mariano, R. S. (2002): "Comparing predictive accuracy". *Journal of Business & Economic Statistics* 20.1, pp. 134–144.

Durbin, J. and Koopman, S. J. (2012): *Time series analysis by state space methods.* Oxford University Press.

D'Innocenzo, E., Lucas, A., Schwaab, B. and Zhang, X. (2024): "Modeling extreme events: Time-varying extreme tail shape". *Journal of Business & Economic Statistics* 42.3, pp. 903–917.

Francq, C. and Zakoïan, J.-M. (2004): "Maximum likelihood estimation of pure GARCH and ARMA-GARCH processes". *Bernoulli* 10.4, pp. 605–637.

Francq, C. and Zakoïan, J.-M. (2010): *GARCH models: Structure, statistical inference and financial applications*. John Wiley & Sons.

Gorton, G. B., Klee, E. C., Ross, C. P., Ross, S. Y. and Vardoulakis, A. P. (2022): *Leverage and stablecoin pegs*. Tech. rep. National Bureau of Economic Research.

Hafner, C. M. (2020): "Testing for bubbles in cryptocurrencies with time-varying volatility". *Journal of Financial Econometrics* 18.2, pp. 233–249.

Harvey, A. and Lange, R.-J. (2017): "Volatility modeling with a generalized t distribution". *Journal of Time Series Analysis* 38.2, pp. 175–190.

Harvey, A. and Lange, R.-J. (2018): "Modeling the interactions between volatility and returns using EGARCH-M". *Journal of Time Series Analysis* 39.6, pp. 909–919.

Harvey, A. and Luati, A. (2014): "Filtering with heavy tails". *Journal of the American Statistical Association* 109.507, pp. 1112–1122.

Harvey, A. and Palumbo, D. (2023): "Score-driven models for realized volatility". *Journal of Econometrics* 237.2, p. 105448.

Harvey, A. C. (2013): *Dynamic models for volatility and heavy tails: with applications to financial and economic time series*. Econometric Society Monographs. Cambridge University Press.

Harvey, A. C. (2022): "Score-driven time series models". *Annual Review of Statistics and Its Application* 9.1, pp. 321–342.

Scaillet, O., Treccani, A. and Trevisan, C. (2020): "High-frequency jump analysis of the bitcoin market". *Journal of Financial Econometrics* 18.2, pp. 209–232.

Zhu, D. and Galbraith, J. W. (2010): "A generalized asymmetric Student-t distribution with application to financial econometrics". *Journal of Econometrics* 157.2, pp. 297–305.

Appendices

A Invertibility of the homoscedastic location model

A.1 generalized t scores

Consider $\frac{\partial u_t}{\partial \mu_{t|t-1}}$ given by

$$\frac{\partial u_t}{\partial \mu_{t|t-1}} = \eta \frac{|\varepsilon_t|^{2\upsilon-2} - \eta(\upsilon-1)\,|\varepsilon_t|^{\upsilon-2}}{(\eta + |\varepsilon_t|^{\upsilon})^2} \tag{32}$$

Let $x = |\varepsilon_t|^{\upsilon} \geq 0$. We have

$$f(x) := \frac{\partial u_t}{\partial \mu_{t|t-1}} = \eta \frac{x^{\frac{2\upsilon-2}{\upsilon}} - \eta(\upsilon-1)x^{\frac{\upsilon-2}{\upsilon}}}{(\eta+x)^2}$$

and

$$\frac{\partial f(x)}{\partial x} = \frac{\frac{\eta}{\upsilon}x^{-\frac{2}{\upsilon}}\left[-2x^2 + \eta(\upsilon+4)(\upsilon-1)x - \eta^2(\upsilon-1)(\upsilon-2)\right]}{(\eta+x)^3} \tag{33}$$

Suppose that $\upsilon > 2$, and denote the two roots of the equation

$$-2x^2 + \eta(\upsilon+4)(\upsilon-1)x - \eta^2(\upsilon-1)(\upsilon-2) = 0$$

by $x^{\pm} = \frac{\eta}{4}\left[(\upsilon+4)(\upsilon-1) \pm \upsilon\sqrt{(\upsilon+7)(\upsilon-1)}\right]$. It is straightforward to show that $\frac{\partial^2 f(x)}{\partial x^2}\big|_{x=x^+} < 0$ and $\frac{\partial^2 f(x)}{\partial x^2}\big|_{x=x^-} > 0$, hence $f(x)$ attains its maximum at x^+ and its minimum at x^-. Therefore, the sufficient condition becomes

$$\max\{\left|\phi + \kappa\eta f(x^-)\right|, \left|\phi + \kappa\eta f(x^+)\right|\} < 1.$$

For the case $\nu = 2$, i.e. the standard Student t distribution, $f(x)$ simplifies to $f(x) = (x-\eta)/(\eta+x)^2$, which takes its maximum value of $1/8\eta$ at $x = 3\eta$, and its minimum value of $-1/\eta$ at $x = 0$, so that the condition reduces to

$$\max\{|\phi - \kappa|, |\phi + \frac{\kappa}{8}|\} < 1.$$

Finally, note that in the case of $1 \leq \upsilon < 2$, $\lim_{x\to 0} f(x) = -\infty$, so the sufficient condition is not satisfied.

A.2 QSD model with EGB2 scores

The score is $u_t = \frac{2\xi+1}{h\xi}\sigma\frac{\exp\left(\left(y_t-\mu_{t|t-1}\right)h/\sigma\right)-1}{\exp\left(\left(y_t-\mu_{t|t-1}\right)h/\sigma\right)+1} = \frac{2\xi+1}{h\xi}\sigma(2b_t-1)$. It is straightforward to show that

$$\frac{\partial u_t}{\partial \mu_{t|t-1}} = -2\frac{2\xi+1}{\xi}b_t(1-b_t) \leq 0 \tag{34}$$

Furthermore, as $b_t = \frac{\exp\left(\left(y_t-\mu_{t|t-1}\right)h/\sigma\right)}{\exp\left(\left(y_t-\mu_{t|t-1}\right)h/\sigma\right)+1} \in (0,1)$, we have $0 < b_t(1-b_t) \leq \frac{1}{4}$, thus the minimum value of $\frac{\partial u_t}{\partial \mu_{t|t-1}}$ is given by $-\frac{2\xi+1}{2\xi}$. Therefore, the sufficient becomes

$$\max(\phi, |\phi - \kappa(1+1/2\xi)|) < 1. \tag{35}$$

For the special case of a logistic distribution, i.e. $\xi = 1$, this condition simplifies to $\max(\phi, \left|\phi - \frac{3}{2}\kappa\right|) < 1$.

B Invertibility of the homoscedastic bubble and balloon model

For the homoscedastic model, the location score from a EGB2 distribution is

$$u_t = \frac{2\xi+1}{h\xi}\sigma\frac{\exp\left(\left(y_t-\mu_{t|t-1}\right)h/\sigma\right)}{\exp\left(\left(y_t-\mu_{t|t-1}\right)h/\sigma\right)+1}.$$

The score u_t is bounded by $\bar{u} = \frac{2\xi+1}{h\xi}\sigma < \infty$, and hence β_t is bounded by $\bar{\beta} = \frac{\kappa}{1-\phi}\bar{u}$. Additionally, we have

$$\frac{\partial \mu_{t|t-1}}{\partial \beta_{t|t-1}} = \alpha_1 \exp(\beta_{t|t-1}) + \alpha_2 \exp(-\beta_{t|t-1}),$$

which is bounded by $\bar{\alpha}(\exp(\bar{\beta}) + \exp(-\bar{\beta}))$, where $\bar{\alpha} = \max(\alpha_1, \alpha_2)$. We now consider $\frac{\partial u_t}{\partial \mu_{t|t-1}}$. It is straightforward to show that

$$\frac{\partial u_t}{\partial \mu_{t|t-1}} = -2\frac{2\xi+1}{\xi} b_t(1-b_t) < 0$$

Furthermore, as $b_t = \frac{\exp\Big((y_t - \mu_{t|t-1})h/\sigma\Big)}{\exp\Big((y_t - \mu_{t|t-1})h/\sigma\Big)+1} \in (0,1)$, we have $0 < b_t(1-b_t) \leq \frac{1}{4}$, thus the minimum value of $\frac{\partial u_t}{\partial \mu_{t|t-1}}$ is given by $-\frac{2\xi+1}{2\xi}$. Therefore, the sufficient condition for invertibility becomes

$$\max(\phi, |\phi - \kappa\overline{\alpha}(\exp(\bar{\beta}) + \exp(-\bar{\beta}))(1 + 1/2\xi)| < 1. \tag{36}$$

C Feasible invertibility conditions

We plot in Figure 11 the filtered time-varying $\beta_{t|t-1}$ and conditional mean with $\beta_{t|t-1}$ initialized with different starting values. The parameters used are based on estimates from fitting the QSD-balloon model to the Bitcoin return series in our empirical analysis. The lines in the plots correspond to the initial value of $\beta_0 + c \cdot \sigma(\beta)$, with $\beta_0 = 0$ being the unconditional mean, $c = (-2, -1.8, \cdots, 0, \cdots, +1.8, +2)$ and $\sigma(\hat{\beta})$ is the standard deviations of the filtered $\beta_{t|t-1}$. We can see that the $\beta_{t|t-1}$ and thus the conditional mean converge to the same trajectory very quickly, suggesting that the filter is invertible at the estimates obtained in our application.

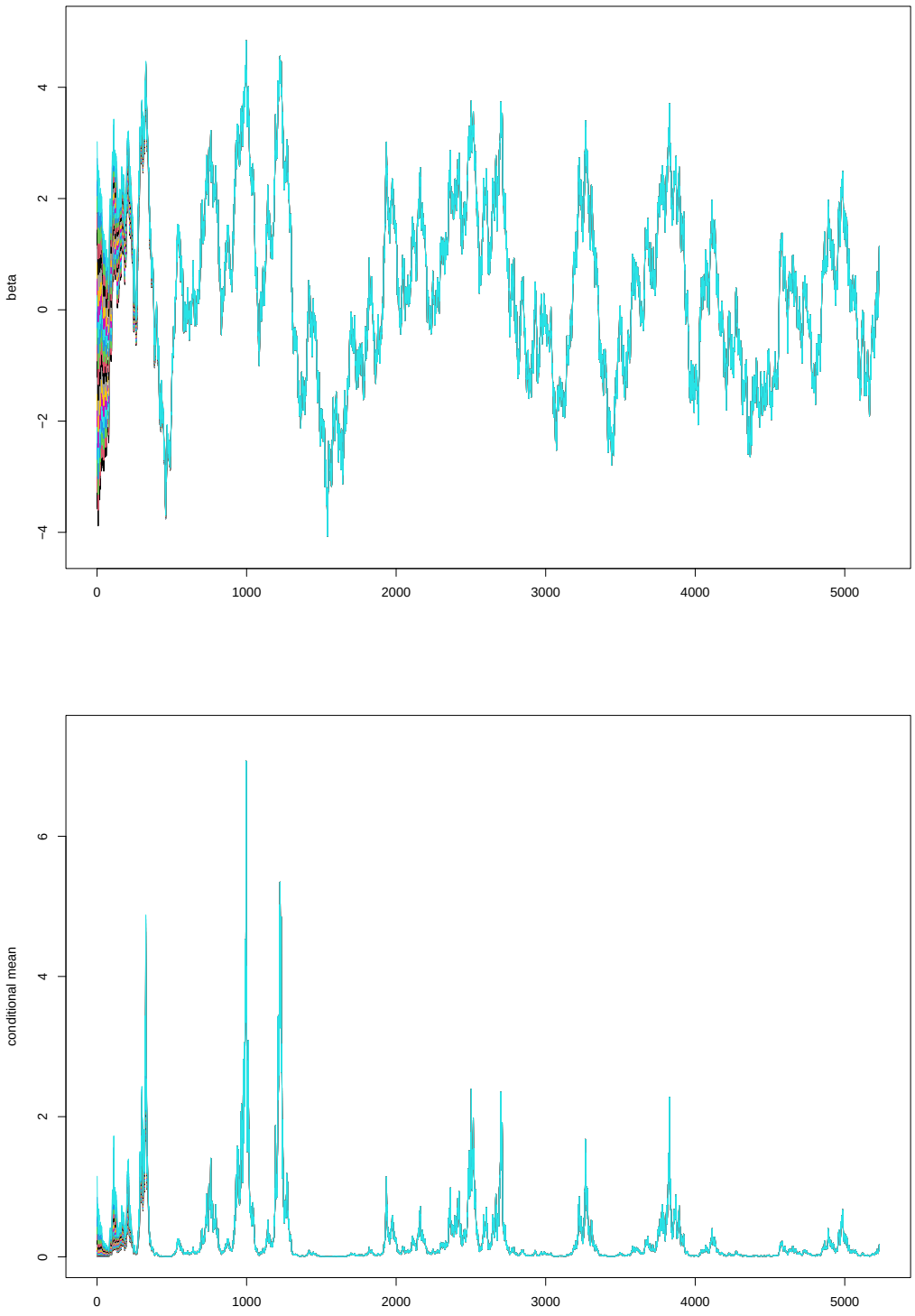


Figure 11: Filtered paths of $\beta_{t|t-1}$ and conditional mean with $\beta_{t|t-1}$ initialized with different starting values.