

A latent-factor-driven endogenous regime-switching non-Gaussian model: Evidence from simulation and application

Ruijun Bu¹ | Jie Cheng² | Fredj Jawadi³

¹University of Liverpool, Liverpool, UK

²Keele University, Keele, UK

³University of Lille, Lille, France

Correspondence

Ruijun Bu, Management School, Chatham Street, Liverpool, L69 7ZH, UK.

Email: ruijunbu@liv.ac.uk

Abstract

Regime-switching models are widely used in empirical economics and finance research for their ability to identify and account for the impact of latent regimes or states on the behaviour of the interested variables. Meanwhile, empirical evidence often reveals complex non-Gaussian characteristics in the state-dependent dynamics. We generalize the latent-factor-driven endogenous regime-switching Gaussian model of Chang, et al., *Journal of Econometrics*, 2017, **196**, 127–143 by allowing the state-dependent conditional distributions to be non-Gaussian. Our setup is more general and promises substantially broader relevance and applicability to empirical studies. We provide evidence to justify our generalization by a simulation study and a real data application. Our simulation results confirm that when the state-dependent dynamics are misspecified, the bias of model parameter estimates, the power of the likelihood ratio test against endogenous regime changes, and the quality of the extracted latent factor all deteriorate quite considerably. In addition, our application to the S&P 500 index return data reveals strong evidence in favour of our non-Gaussian assumption, and the superiority of our model specification delivers an important risk management implication.

KEYWORDS

endogeneity, fat tails, latent factor, non-Gaussian process, regime switching

1 | INTRODUCTION

Since the seminal work of Hamilton (1988, 1989), regime-switching time series models have been

extensively used in empirical economics and finance research. These models assume that there is some unobserved latent factor implicitly driving the evolution of changing regimes or states, representing varied economic or market conditions under which the behaviours of the concerned economic or financial variables are investigated. Their intrinsic ability to identify and account for the impact of changing regimes on the behaviour of the time series variables makes regime-switching

We thank the editor and two anonymous referees for many helpful comments. We also thank the organizers and participants of The 4th International Workshop in Financial Markets and Nonlinear Dynamics (FMND, www.fmnd.fr).

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2020 The Authors. International Journal of Finance & Economics published by John Wiley & Sons Ltd.

models a powerful and attractive econometric tool for many empirical applications. See, for example, Hamilton (1988, 1989), Kim (1994), Garcia and Perron (1996), Bollen, Gray, and Whaley (2000), among others.

Not surprisingly, most of the early studies on regime-switching models were established under the classic linear-Gaussian framework. In recent years, however, more and more researchers have introduced regime-switching mechanism to more general classes of non-linear and non-Gaussian models. Most obvious examples of these models include, but are not limited to, regime-switching models with fat-tailed error distributions, regime-switching duration models, regime-switching continuous-time diffusion models and so on. See Choi (2009), Goutte and Zou (2013) and Bu, Cheng, and Hadri (2017) for some recent examples. Conventional regime-switching models assume that the transition of the underlying finite state Markov chain is independent of the evolutionary path of the observed time series variable, which often proves to be unrealistic in many cases. For example, studies including Diebold, Lee, and Weinbach (1994), Kim, Piger, and Startz (2008), Choi (2009), Bu et al. (2017), Chang, Choi, and Park (2017) and others all reported evidence of endogenous regime changes that we observe widely and frequently for many practical applications.

There are several different ways of modelling endogenous regime changes. For example, Choi (2009) and Bu et al. (2017) chose to model the state transition probabilities explicitly as a function of the lagged values of the observed variable. Diebold et al. (1994) considered a Markov process that is driven by a set of observed variables. Kim et al. (2008) studied a regime-switching model driven by an endogenous identically and independently distributed latent factor with the threshold level determined by the previous state and possibly lagged values of the time series. Most recently, Chang et al. (2017) (CCP) allows the regime changes to be determined by an autoregressive latent factor assumed to be correlated with the previous innovation of the state-dependent process. By construction, the resulting state transition probabilities are time-varying and dependent on the lagged values of the observed time series. Their approach also allows the users to extract the unobserved latent factor, which can subsequently be analysed in conjunction with observable economic variables to better understand the dynamics of economic trends.

It is important to note that a crucial assumption underlying the endogenous regime-switching framework of CCP is the conditional normality of the state-dependent processes. Although this assumption is sufficient for some empirical studies, it may turn out to be too restrictive and unrealistic for many other situations. For

example, many financial time series exhibit heavy-tail characteristics, and for certain data type, for example, the duration data, the state-dependent conditional distributions must have non-negative support. In addition, there is a growing literature on regime-switching continuous-time processes, which, except for the Ornstein Uhlenbeck process, generally have non-Gaussian transition densities. Hence, the applicability of CCP's original setup, despite its novelty in terms of model interpretability and the ability to extract the latent factor, can be quite limited without some necessary extensions.

For this reason, we propose in this paper a generalized latent factor-driven endogenous regime-switching model where the state-dependent processes are allowed to be non-Gaussian. Our approach is built upon the framework of CCP, who considered a regime-switching Gaussian model driven by whether a latent factor is above or below an unknown threshold, assuming that the innovations of the state-dependent process and the latent process are bivariate normal. We extend their model by allowing the conditional distributions of the state-dependent processes to be non-Gaussian. Clearly, our setting lends itself to the applications in much wider and more realistic situations, especially in the modelling of highly complicated financial and economic dynamics.

To accommodate non-Gaussian state-dependent conditional distributions, we consider a simple distribution transformation strategy. Specifically, conditional non-Gaussian state-dependent random variables are first transformed, in a one-to-one sense, into standard normal random innovations by a distribution transformation. We then assume, as in CCP, that the resulting standard normal random innovation and the innovation of the latent factor are jointly normal. This setup, the nonlinear transformation and the dependence structure between the two innovations, allows us to deviate from the restrictive normality assumption and at the same time allows us to adopt the inferential procedures of CCP relatively easily with some simple modifications. Following closely the results of CCP, we develop a modified Markov filter for our system, which leads to a simple Maximum Likelihood (ML) estimation procedure, and a modified procedure for extracting the latent factor. The consistency and asymptotic normality of the ML estimator for general regime-switching models follow from the recent work by Kasahara and Shimotsu (2019). It is not difficult to see that our setup and inferential procedures reduce to the case of CCP when the state-dependent processes are linear with Gaussian innovations. Just as CCP, our approach is equally intuitive and easy to implement, but it is more general and more applicable.

To demonstrate the importance of the ability to deviate from the Gaussian assumption on the state-dependent

conditional distributions and the consequence of potential misspecifications, we conduct a simulation exercise based on the same regime-switching volatility model studied by CCP. Our focus is on the impact of misspecification by the normal assumption on the bias of the model parameter estimates, the power of the Likelihood Ratio (LR) test against endogenous regime changes, and the quality in term of Average Mean Squared Error (AMSE) of the extracted latent factor, when the true state-dependent conditional distributions are non-Gaussian. We assume that the true data-generating error distribution is the Student- t distribution and examine the difference in the above statistics between the correctly specified non-Gaussian models and the incorrectly specified Gaussian models. Unsurprisingly, we find that such misspecifications potentially have negative impacts on all of these criteria. Specifically, across all the scenarios we considered in our simulation, such misspecifications led to increased bias on all parameter estimates, reduced power of the LR test, and increased AMSE of the extracted latent factor.

We also demonstrate the importance of non-Gaussian assumptions through our application to the daily S&P 500 index return data. Specifically, we considered the same regime-switching volatility model as in our simulation with alternative error distributions, namely the normal distribution and the Student- t distribution. We find strong evidence in support of endogenous regime changes, regardless of error specifications. More importantly, our non-Gaussian models outperform the Gaussian models statistically significantly in the sample in terms of the log-likelihood and out of the sample in terms Mean Squared Predictive Error (MSPE) and Gaussian Quasi-Likelihood (QLIKE) for volatility forecasting. To investigate the potential economic implication of making appropriate distributional assumptions on the state-dependent dynamics in our models, we examine the difference between the two error specifications in a financial risk management application. Since our focus is on the fat-tailedness in financial returns, we compare our models in terms of Conditional Tail Expectation (CTE) forecasts. We find that our fat-tailed Student- t model always produces statistically more extreme CTE forecasts than does the Gaussian model. This means that the model with no fat-tailed error distribution significantly underestimated the tail risks relative to our model with the more adequate fat-tailed Student- t error specification and that such under-estimation can be quite substantial particularly during crisis periods where extreme movements of the market are observed more frequently.

The rest of this paper is organized as follows: In Section 2, we briefly review the latent factor-driven

regime-switching model of CCP and then generalize it to accommodate general non-Gaussian state-dependent processes. Section 3 presents our simulation study, focusing on the impacts of misspecification by the normal assumption on the estimation and inference of non-Gaussian models. In Section 4, we further demonstrate the importance of our extension by presenting an empirical application to the S&P 500 index return series. Some concluding remarks are provided in Section 5.

2 | LATENT-FACTOR-DRIVEN ENDOGENOUS REGIME-SWITCHING MODELS

In this section, we briefly outline the framework developed by CCP for modelling endogenous regime-switching conditional Gaussian processes. We then extend it to the case where the state-dependent processes are general, that is, non-Gaussian, and discuss model estimation and inference of the latent factor under the new framework.

2.1 | Conditional Gaussian state-dependent processes

Conventional two-state regime-switching models typically assume that the evolution of the unobserved random state process (s_t), taking value 1 or 0 representing high or low state of the economy, is governed by a stationary two-state Markov chain with state transition probability matrix

$$M = \begin{bmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{bmatrix} = \begin{bmatrix} p_{00} & 1-p_{00} \\ 1-p_{11} & p_{11} \end{bmatrix},$$

where

$$p_{00} = \mathbb{P}\{s_t = 0 | s_{t-1} = 0\}$$

$$p_{11} = \mathbb{P}\{s_t = 1 | s_{t-1} = 1\}$$

are the probabilities of the Markov chain remaining in the same state from $t-1$ to t .

Let y_t be the observed time series variable we wish to model. Then, the dynamics of y_t conditional on $t-1$ and s_t is assumed to be governed by a state-dependent process with unknown parameter vector $\bar{\pi}$ and $\underline{\pi}$ for high and low states, respectively. For notational brevity, we define the following two specific filtrations for time series (y_t) and state process (s_t):

$$\mathcal{Y}_{t-a,t-b} = \{y_{t-a}, y_{t-a-1}, \dots, y_{t-b}\},$$

$$\mathcal{S}_{t-a,t-b} = \{s_{t-a}, s_{t-a-1}, \dots, s_{t-b}\},$$

for $0 \leq a < b \leq t-1$. The usual filtration of y_t is then denoted by $\mathcal{F}_{t-1} = \mathcal{Y}_{t-1,1} = \{y_{t-1}, y_{t-2}, \dots, y_1\}$.

If the evolution of s_t is independent of $\mathcal{Y}_{t-1,1}$, then p_{00} and p_{11} are constant and the switching of regimes is exogenous. However, this assumption often needs to be relaxed to enrich the dynamics of the model by introducing endogenous regime switches. Studies including Choi (2009) and Bu et al. (2017), for instance, specify p_{00} and p_{11} as a function of $\mathcal{Y}_{t-1,1}$, so that the probabilities of regime changes are time-varying and endogenously influenced by the past realizations of y_t . The advantage of modelling endogeneity in this way is that one can directly and quite flexibly specify the time-varying transition probabilities as a function of lagged values of y_t . The disadvantage, however, is that such endogenous dependence is usually arbitrary and often does not provide much economic intuition or justification as to what is driving the switching of regimes.

To provide an alternative solution, CCP proposed a novel and more intuitive approach for modelling endogenous regime changes. They assume that the evolution of the state variable s_t is determined by whether an unobserved latent factor (w_t) is above/equal or below some unknown threshold level τ . Formally, they define

$$s_t = 1\{w_t \geq \tau\} \quad (1)$$

where $1\{\cdot\}$ is the indicator function and (w_t) is a zero-mean AR(1) process

$$w_t = \alpha w_{t-1} + v_t \quad (2)$$

with parameter $\alpha \in (-1, 1]$ and i.i.d. standard normal innovations (v_t). The advantage of their setup is that endogenous regime changes can be introduced by allowing v_{t+1} to be correlated with the innovation term u_t of their state-dependent process. More specifically, they assume that their state-dependent dynamics is governed by the following conditional Gaussian process

$$y_t = m(\mathcal{S}_{t,t-k}, \mathcal{Y}_{t-1,t-k}; \vartheta) + \sigma(\mathcal{S}_{t,t-k}, \mathcal{Y}_{t-1,t-k}; \vartheta) u_t \quad (3)$$

where $m(\mathcal{S}_{t,t-k}, \mathcal{Y}_{t-1,t-k}; \vartheta)$ and $\sigma(\mathcal{S}_{t,t-k}, \mathcal{Y}_{t-1,t-k}; \vartheta)$ are the mean and the volatility functions, respectively, with $\vartheta = (1-s_t)\underline{\mu} + s_t\bar{\mu}$. Most importantly, u_t and v_{t+1} are jointly i.i.d. as

$$\begin{pmatrix} u_t \\ v_{t+1} \end{pmatrix} = d\mathbb{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\right) \quad (4)$$

with unknown correlation coefficient parameter ρ .

In this framework, α controls the persistency of regime changes. When $|\alpha| < 1$, the latent factor (w_t) is asymptotically stationary, and it becomes strictly stationary if in addition $w_0 = d\mathbb{N}(0, 1/(1 - \alpha^2))$. This then implies that the state process s_t is also stationary and the unconditional state probabilities exist. When $\alpha = 1$, however, (w_t) becomes a unit root process and its transition density is a function of time. Consequently, the state process s_t becomes non-stationary and the unconditional state probabilities do not exist. On the other hand, the parameter ρ in the joint distribution (4) determines the level of endogeneity of regime changes. If $\rho \neq 0$, the state transition probabilities become time-varying and dependent on the lagged values of y_t . As ρ approaches unity in modulus, the endogeneity of regime changes driven by (w_t) becomes stronger. When $|\rho| = 1$, there is perfect endogeneity, and consequently the current shock u_t fully dictates the realization of the latent factor w_{t+1} determining the state in the next period. In this case, the state process is actually degenerate and becomes adapted to \mathcal{F}_{t-1} .

CCP derived the endogenous state transition probabilities of the state process s_t and developed a modified Markov switching filter to facilitate the ML estimation of their model. They also derived the endogenous transition density of the latent factor w_t and developed a procedure for inferring the latent factor conditional on \mathcal{F}_t . Readers can refer to Section 3 of their paper for a complete understanding. In the next section, we propose a generalization of their framework and provide more general results. As a special case, their results are readily available from ours by imposing the restriction that the state-dependent process is conditional Gaussian.

2.2 | General state-dependent processes

2.2.1 | The setup

The conditional Gaussian assumption is quite crucial in CCP's setup, because it allows them to model the joint distribution of (u_t, v_{t+1}) as the bivariate standard normal distribution, which is convenient for their subsequent technical analysis. However, as discussed above, the conditional Gaussian assumption on the state-dependent process is clearly restrictive, preventing wider applications of their modelling strategy to many realistic situations. To overcome this limitation, we propose a relatively simple generalization of their framework to accommodate general state-dependent processes.

We assume that the state process s_t and the latent factor w_t are still defined by (1) and (2), respectively, as in CCP. Crucially, instead of assuming a conditional Gaussian state-dependent process in (3), we define our state-dependent process quite flexibly as a general conditional stochastic process, which we denote as

$$y_t = \{y_t | \mathcal{S}_{t,t-k}, \mathcal{Y}_{t-1,t-k}, \vartheta = (1-s_t)\underline{\mu} + s_t\bar{\mu}\} \quad (5)$$

for some known value of k and $t = 1, 2, \dots$. Moreover, we assume that the probability law of y_t conditional on $(\mathcal{S}_{t,t-k}, \mathcal{Y}_{t-1,t-k}, \vartheta)$ is governed by its conditional cumulative distribution function (cdf) $F(y_t | \mathcal{S}_{t,t-k}, \mathcal{Y}_{t-1,t-k}; \vartheta)$ with continuous probability density function (pdf) $f(y_t | \mathcal{S}_{t,t-k}, \mathcal{Y}_{t-1,t-k}; \vartheta)$. Most importantly, for all t we define the standard normal cdf inverse transform of the conditional cdf of y_t as

$$\tilde{u}_t = \Phi^{-1}(F(y_t | \mathcal{S}_{t,t-k}, \mathcal{Y}_{t-1,t-k}; \vartheta)) \quad (6)$$

where $\Phi(\cdot)$ is the standard normal cdf function. It is important to note that since y_t is Markovian with respect to $\mathcal{S}_{t,t-k}$ and $\mathcal{Y}_{t-1,t-k}$, \tilde{u}_t is i.i.d standard normal random variable by construction. Consequently, we can further assume, in the spirit of CCP, that

$$\begin{pmatrix} \tilde{u}_t \\ v_{t+1} \end{pmatrix} = d\mathbb{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\right), \quad (7)$$

which gives rise to the possibility of endogenous regime switches. This completes the setup of the generalized framework.

We can see that the inverse cdf transform (6) is the key to our framework. It allows us to be freed from the conditional Gaussian restriction, while at the same time allows us to continue to take advantages of the analytical tractability of the bivariate normal distribution in (7), which links the innovations of state-dependent process y_t and latent factor w_{t+1} . Our setup clearly encompasses the framework of CCP as a special case when the state-dependent process is conditional Gaussian. This is because

$$\begin{aligned} \tilde{u}_t &= \Phi^{-1}(F(y_t | \mathcal{S}_{t,t-k}, \mathcal{Y}_{t-1,t-k}; \vartheta)) \\ &= \Phi^{-1}\left(\Phi\left(\frac{y_t - m(\mathcal{S}_{t,t-k}, \mathcal{Y}_{t-1,t-k}; \vartheta)}{\sigma(\mathcal{S}_{t,t-k}, \mathcal{Y}_{t-1,t-k}; \vartheta)}\right)\right) = u_t. \end{aligned} \quad (8)$$

In fact, our analytical results provided in the next section can be quite easily derived by following the logic in their paper. In the meantime, their results will immediately follow if (8) holds.

2.2.2 | ML estimation and extraction of latent factor

The ML estimation of the generalized model can proceed in a similar manner as in CCP. Given n observations $\{y_1, y_2, \dots, y_n\}$ of the time series y_t , the log-likelihood function of the model can be written as

$$\ell(y_1, \dots, y_n) = \log p(y_1) + \sum_{t=2}^n \log p(y_t | \mathcal{F}_{t-1}) \quad (9)$$

Of course, the log-likelihood function includes the vector of unknown parameters $\theta \in \Theta$. It is, however, suppressed for notational brevity. As usual, the ML estimator $\hat{\theta}$ of θ is given by

$$\hat{\theta} = \arg \max_{\theta \in \Theta} \ell(y_1, \dots, y_n)$$

For the model given by (1), (2), and (5)–(7), θ consists of the state-dependent parameter ϑ , the autoregressive coefficient α of the latent factor, the threshold level τ , as well as the endogeneity parameter ρ .

As with most econometric models containing latent components, ML estimation of our model requires some filtering procedure to evaluate the conditional density function $p(y_t | \mathcal{F}_{t-1})$ in (9). As discussed by CCP, the conventional Markov switching filter is no longer applicable, since the state process (s_t) defined in (1) is not a Markov chain unless $\rho = 0$. We consider the same modified Markov switching filter as in CCP, which consists of the usual prediction and updating steps. Specifically, the required conditional density function $p(y_t | \mathcal{F}_{t-1})$ can be written as

$$p(y_t | \mathcal{F}_{t-1}) = \sum_{s_t} \cdots \sum_{s_{t-k}} p(y_t, \mathcal{S}_{t,t-k} | \mathcal{F}_{t-1}) p(\mathcal{S}_{t,t-k} | \mathcal{F}_{t-1}).$$

For the prediction step, we note that

$$\begin{aligned} p(\mathcal{S}_{t,t-k} | \mathcal{F}_{t-1}) &= \sum_{s_{t-k-1}} p(s_t | \mathcal{S}_{t-1,t-k-1}, \mathcal{F}_{t-1}) p(\mathcal{S}_{t-1,t-k-1} | \mathcal{F}_{t-1}) \end{aligned}$$

and for the updating step, we have

$$\begin{aligned} p(\mathcal{S}_{t,t-k} | \mathcal{F}_t) &= p(\mathcal{S}_{t,t-k} | y_t, \mathcal{F}_{t-1}) \\ &= \frac{p(y_t | \mathcal{S}_{t,t-k}, \mathcal{F}_{t-1}) p(\mathcal{S}_{t,t-k} | \mathcal{F}_{t-1})}{p(y_t | \mathcal{F}_{t-1})} \end{aligned}$$

Note that $p(y_t | \mathcal{S}_{t,t-k}, \mathcal{F}_{t-1}) = f(y_t | \mathcal{S}_{t,t-k}, \mathcal{Y}_{t-1,t-k}; \vartheta)$ is known from the parametric specification of the

state-dependent process. Therefore, the knowledge of $p(s_t | \mathcal{S}_{t-1,t-k-1}, \mathcal{F}_{t-1}) = p(s_t | \mathcal{S}_{t-1,t-k-1}, \mathcal{Y}_{t-1,t-k-1})$, which is the endogenous state transition probability, will be sufficient for the implementation of the filter. Following closely the logic behind the results in CCP, we now provide details of the required endogenous state transition probability and summarize them as follows:

Theorem 2.1 *Let $|\rho| < 1$. The bivariate process (s_t, y_t) on $\{0, 1\} \times \mathbb{R}$ is a $(k+1)$ th order Markov process, whose transition density with respect to the product of the counting and Lebesgue measure is given by*

$$\begin{aligned} p(s_t, y_t | \mathcal{S}_{t-1,t-k-1}, \mathcal{Y}_{t-1,t-k-1}) \\ = p(y_t | \mathcal{S}_{t,t-k}, \mathcal{Y}_{t-1,t-k}) p(s_t | \mathcal{S}_{t-1,t-k-1}, \mathcal{Y}_{t-1,t-k-1}) \end{aligned}$$

where $p(y_t | \mathcal{S}_{t,t-k}, \mathcal{Y}_{t-1,t-k})$ is the state-dependent transition density of y_t and

$$\begin{aligned} p(s_t | \mathcal{S}_{t-1,t-k-1}, \mathcal{Y}_{t-1,t-k-1}) \\ = (1-s_t) \omega_\rho(\mathcal{S}_{t-1,t-k-1}, \mathcal{Y}_{t-1,t-k-1}) \\ + s_t [1 - \omega_\rho(\mathcal{S}_{t-1,t-k-1}, \mathcal{Y}_{t-1,t-k-1})] \end{aligned}$$

with transition probability ω_ρ of the endogenous state process (s_t) to low state ($s_t = 0$). Let $\Phi_p(x) = \Phi(x/\sqrt{1-\rho^2})$. If $|\alpha| < 1$, ω_ρ is given by

$$\begin{aligned} \omega_\rho(\mathcal{S}_{t-1,t-k-1}, \mathcal{Y}_{t-1,t-k-1}) \\ = \frac{\left[(1-s_{t-1}) \int_{-\infty}^{\tau\sqrt{1-\alpha^2}} + s_{t-1} \int_{\tau\sqrt{1-\alpha^2}}^{\infty} \right] \Phi_\rho\left(\tau - \rho\tilde{u}_{t-1} - \frac{\alpha x}{\sqrt{1-\alpha^2}}\right) \varphi(x) dx}{(1-s_{t-1}) \Phi(\tau\sqrt{1-\alpha^2}) + s_{t-1} [1 - \Phi(\tau\sqrt{1-\alpha^2})]}, \end{aligned}$$

and, if $\alpha = 1$, for $t = 1$, $\omega_\rho(s_0) = \Phi(\tau)$ with $\Pr\{s_0 = 0\} = 1$ and $\Pr\{s_0 = 1\} = 1$ respectively when $\tau > 0$ and $\tau \leq 0$ and, for $t \geq 2$,

$$\begin{aligned} \omega_\rho(\mathcal{S}_{t-1,t-k-1}, \mathcal{Y}_{t-1,t-k-1}) \\ = \frac{\left[(1-s_{t-1}) \int_{-\infty}^{\tau/\sqrt{t-1}} + s_{t-1} \int_{\tau/\sqrt{t-1}}^{\infty} \right] \Phi_\rho(\tau - \rho\tilde{u}_{t-1} - x\sqrt{t-1}) \varphi(x) dx}{(1-s_{t-1}) \Phi(\tau/\sqrt{t-1}) + s_{t-1} [1 - \Phi(\tau/\sqrt{t-1})]}. \end{aligned}$$

Theorem 2.1 fully specifies the joint transition of (s_t) and (y_t) for the case $|\rho| < 1$, and Corollary 2.1 below gives explicit details for the case $|\rho| = 1$.

Corollary 2.1 *If $|\rho| = 1$, the transition probability ω_ρ of the endogenous state process (s_t) to low state ($s_t = 0$)*

conditional on previous states and past observed time series is given as follows: (a) If $\alpha = 0$,

$$\omega_\rho(\mathcal{S}_{t-1,t-k-1}, \mathcal{Y}_{t-1,t-k-1}) = \begin{cases} 1 & \text{if } \rho\tilde{u}_{t-1} < \tau \\ 0 & \text{otherwise} \end{cases}$$

(b) If $0 < \alpha < 1$,

$$\begin{aligned} \omega_\rho(\mathcal{S}_{t-1,t-k-1}, \mathcal{Y}_{t-1,t-k-1}) \\ = (1-s_{t-1}) \min\left(1, \frac{\Phi\left((\tau - \rho\tilde{u}_{t-1}) \frac{\sqrt{1-\alpha^2}}{\alpha}\right)}{\Phi(\tau\sqrt{1-\alpha^2})}\right) \\ + s_{t-1} \max\left(0, \frac{\Phi\left((\tau - \rho\tilde{u}_{t-1}) \frac{\sqrt{1-\alpha^2}}{\alpha}\right) - \Phi(\tau\sqrt{1-\alpha^2})}{1 - \Phi(\tau\sqrt{1-\alpha^2})}\right) \end{aligned}$$

(c) If $-1 < \alpha < 0$,

$$\begin{aligned} \omega_\rho(\mathcal{S}_{t-1,t-k-1}, \mathcal{Y}_{t-1,t-k-1}) \\ = s_{t-1} \min\left(1, \frac{1 - \Phi\left((\tau - \rho\tilde{u}_{t-1}) \frac{\sqrt{1-\alpha^2}}{\alpha}\right)}{1 - \Phi(\tau\sqrt{1-\alpha^2})}\right) \\ + (1-s_{t-1}) \max\left(0, \frac{\Phi(\tau\sqrt{1-\alpha^2}) - \Phi\left((\tau - \rho\tilde{u}_{t-1}) \frac{\sqrt{1-\alpha^2}}{\alpha}\right)}{\Phi(\tau\sqrt{1-\alpha^2})}\right) \end{aligned}$$

(d) If $\alpha = 1$, for $t = 1$, $\omega_\rho(s_0, y_0) = \Phi(\tau - \tilde{u}_0)$ with $\mathbb{P}\{s_0 = 0\} = 1$ and $\mathbb{P}\{s_0 = 1\} = 1$ respectively when $\tau > 0$ and $\tau \leq 0$ and, for $t \geq 2$,

$$\begin{aligned} \omega_\rho(\mathcal{S}_{t-1,t-k-1}, \mathcal{Y}_{t-1,t-k-1}) \\ = \begin{cases} 1-s_{t-1} & \text{if } \rho\tilde{u}_{t-1} > 0 \\ \frac{\Phi\left((\tau - \rho\tilde{u}_{t-1}) \frac{1}{\sqrt{t-1}}\right) - s_{t-1} \Phi(\tau/\sqrt{t-1})}{(1-s_{t-1}) \Phi(\tau/\sqrt{t-1}) + s_{t-1} [1 - \Phi(\tau/\sqrt{t-1})]} & \text{otherwise.} \end{cases} \end{aligned}$$

In addition to ML estimation of the model, CCP developed a procedure for inferring the latent factor w_t for their model. A similar but slightly modified procedure can be developed for our general framework. Based on the modified filter above for the state process s_t , we can extract the latent factor through the prediction and updating steps described above. In the prediction step, we note that

$$p(w_t, \mathcal{S}_{t-1,t-k} | \mathcal{F}_{t-1}) = \sum_{\mathcal{S}_{t-k-1}} p(w_t | \mathcal{S}_{t-1,t-k-1}, \mathcal{F}_{t-1}) p(\mathcal{S}_{t-1,t-k-1} | \mathcal{F}_{t-1})$$

Then, we may obtain

$$p(w_t, \mathcal{S}_{t-1,t-k} | \mathcal{F}_t) = \frac{p(y_t | w_t, \mathcal{S}_{t-1,t-k}, \mathcal{F}_{t-1}) p(w_t, \mathcal{S}_{t-1,t-k} | \mathcal{F}_{t-1})}{p(y_t | \mathcal{F}_{t-1})}$$

in the updating step. By marginalizing $p(w_t, \mathcal{S}_{t-1,t-k} | \mathcal{F}_t)$ we can then get

$$p(w_t | \mathcal{F}_t) = \sum_{\mathcal{S}_{t-1}} \cdots \sum_{\mathcal{S}_{t-k}} p(w_t, \mathcal{S}_{t-1,t-k} | \mathcal{F}_t) = \sum_{\mathcal{S}_{t-1,t-k}} p(w_t, \mathcal{S}_{t-1,t-k} | \mathcal{F}_t)$$

which yields the inferred latent factor as the following filtered latent factor

$$\mathbb{E}(w_t | \mathcal{F}_t) = \int w_t p(w_t | \mathcal{F}_t) dw_t \quad (10)$$

Therefore, we may easily extract the inferred factor, once the ML estimates of $p(w_t | \mathcal{F}_t)$, $1 \leq t \leq n$, are available. Note that the knowledge of $p(w_t | \mathcal{S}_{t-1,t-k-1}, \mathcal{Y}_{t-1,t-k-1})$ will suffice for the implementation of the procedure. The Corollary 2.2 below gives all the required details.

Corollary 2.2 *The transition density of latent factor (w_t) on previous states and past observed time series is given as follows:*

(a) When $|\alpha| < 1$ and $|\rho| < 1$,

$$p(w_t | s_{t-1} = 1, \mathcal{S}_{t-2,t-k-1}, \mathcal{Y}_{t-1,t-k-1}) = \frac{\left(1 - \Phi\left(\sqrt{\frac{1-\rho^2+\alpha^2\rho^2}{1-\rho^2}}\left(\tau - \frac{\alpha(w_t - \rho\tilde{u}_{t-1})}{1-\rho^2+\alpha^2\rho^2}\right)\right)\right)}{1 - \Phi(\tau\sqrt{1-\alpha^2})} \mathbb{N}\left(\rho\tilde{u}_{t-1}, \frac{1-\rho^2+\alpha^2\rho^2}{1-\rho^2}\right),$$

$$p(w_t | s_{t-1} = 0, \mathcal{S}_{t-2,t-k-1}, \mathcal{Y}_{t-1,t-k-1})$$

$$= \frac{\Phi\left(\sqrt{\frac{1-\rho^2+\alpha^2\rho^2}{1-\rho^2}}\left(\tau - \frac{\alpha(w_t - \rho\tilde{u}_{t-1})}{1-\rho^2+\alpha^2\rho^2}\right)\right)}{\Phi(\tau\sqrt{1-\alpha^2})} \mathbb{N}\left(\rho\tilde{u}_{t-1}, \frac{1-\rho^2+\alpha^2\rho^2}{1-\rho^2}\right).$$

(b) When $|\alpha| < 1$ and $|\rho| = 1$,

$$p(w_t | s_{t-1} = 1, \mathcal{S}_{t-2,t-k-1}, \mathcal{Y}_{t-1,t-k-1})$$

$$= \begin{cases} \frac{\sqrt{1-\alpha^2}}{\alpha} \phi\left(\frac{w_t - \rho\tilde{u}_{t-1}}{\alpha} \sqrt{1-\alpha^2}\right) & \text{if } w_t \geq \alpha\tau + \rho\tilde{u}_{t-1} \\ 0 & \text{otherwise} \end{cases}$$

$$p(w_t | s_{t-1} = 0, \mathcal{S}_{t-2,t-k-1}, \mathcal{Y}_{t-1,t-k-1})$$

$$= \begin{cases} \frac{\sqrt{1-\alpha^2}}{\alpha} \phi\left(\frac{w_t - \rho\tilde{u}_{t-1}}{\alpha} \sqrt{1-\alpha^2}\right) & \text{if } w_t \leq \alpha\tau + \rho\tilde{u}_{t-1} \\ 0 & \text{otherwise} \end{cases}$$

(c) When $\alpha = 1$ and $|\rho| < 1$,

$$p(w_t | s_{t-1} = 1, \mathcal{S}_{t-2,t-k-1}, \mathcal{Y}_{t-1,t-k-1})$$

$$= \frac{\left(1 - \Phi\left(\sqrt{\frac{t-t\rho^2+\rho^2}{1-\rho^2}}\left(\tau - \frac{w_t - \rho\tilde{u}_{t-1}}{t-t\rho^2+\rho^2}\right)\right)\right)}{1 - \Phi(\tau\sqrt{t-1})} \mathbb{N}\left(\rho\tilde{u}_{t-1}, \frac{t-t\rho^2+\rho^2}{t-1}\right),$$

$$p(w_t | s_{t-1} = 0, \mathcal{S}_{t-2,t-k-1}, \mathcal{Y}_{t-1,t-k-1})$$

$$= \frac{\Phi\left(\sqrt{\frac{t-t\rho^2+\rho^2}{1-\rho^2}}\left(\tau - \frac{w_t - \rho\tilde{u}_{t-1}}{t-t\rho^2+\rho^2}\right)\right)}{1 - \Phi(\tau\sqrt{t-1})} \mathbb{N}\left(\rho\tilde{u}_{t-1}, \frac{t-t\rho^2+\rho^2}{t-1}\right).$$

(d) When $\alpha = 1$ and $|\rho| = 1$,

$$p(w_t | s_{t-1} = 1, \mathcal{S}_{t-2,t-k-1}, \mathcal{Y}_{t-1,t-k-1})$$

$$= \begin{cases} \frac{1}{\sqrt{t-1}} \phi\left(\frac{w_t - \rho\tilde{u}_{t-1}}{t-1}\right) & \text{if } w_t \geq \tau + \rho\tilde{u}_{t-1} \\ 0 & \text{otherwise} \end{cases}$$

$$p(w_t | s_{t-1} = 0, \mathcal{S}_{t-2,t-k-1}, \mathcal{Y}_{t-1,t-k-1})$$

$$= \begin{cases} \frac{1}{\sqrt{t-1}} \phi\left(\frac{w_t - \rho\tilde{u}_{t-1}}{t-1}\right) & \text{if } w_t \leq \tau + \rho\tilde{u}_{t-1} \\ 0 & \text{otherwise} \end{cases}$$

3 | SIMULATION

Model misspecifications generally have negative impacts on parameter estimation and inference of econometric models. In the current context, we are particularly interested in the impacts of misspecification in the state-dependent dynamics on the estimation and inference of the model. Among many potential deviations from the Gaussian assumption, the fat-tailedness is by far the most debated and researched issue in empirical economics and finance. In particular, fat-tailedness is an important characteristic of financial returns. Correct specification of the error distributions in the modelling of financial return series has important implications on the practice of financial forecasting, portfolio selection, risk management, and

so on. See, for example, Hansen (1994), Jondeau and Rockinger (2003), and others. For this reason, our simulation study in this section and our empirical application in the next section will both focus on the impacts of ignored fat-tailedness in the error distribution of the state-dependent process on the estimation and inference of the latent factor-driven endogenous regime-switching model.

3.1 | Simulation design

Since our empirical application mainly focuses on the regime-switching feature in the volatility of the S&P 500 index return series, we consider a regime-switching volatility model in our simulation. Specifically, it is specified as

$$y_t = \sigma(s_t)\varepsilon_t, \quad \sigma(s_t) = (1 - s_t)\sigma_L + s_t\sigma_H, \quad (11)$$

where ε_t is an i.i.d. error process with unit variance. The same volatility model was considered by CCP and we will also use this model in our application. Following CCP, we set $\sigma_L = 0.04$ representing the low volatility state L and $\sigma_H = 0.12$ representing the high volatility state H , and to account for different levels of endogeneity, we allow ρ to vary from 0 to -1 on an equal distance grid.¹ As in CCP, we also consider two pairs of the autoregressive coefficient α of the latent factor and the threshold level τ given by $(\alpha, \tau) = (0.4, 0.5)$, $(0.8, 0.7)$. When $\rho = 0$, the first scenario corresponds to $(p_{LL}, p_{HH}) = (0.75, 0.5)$, and the second scenario corresponds to $(p_{LL}, p_{HH}) = (0.86, 0.72)$, both implying the same unconditional state probabilities of $(2/3, 1/3)$. Up to this point, our setup is identical to that of CCP.

Assuming that ε_t is i.i.d. standard normal, CCP investigated the impact of ignored endogeneity on the bias in the estimation of model parameters. Distinct from theirs, the focus of our simulation is on the impact of misspecified state-dependent dynamics. To achieve this, we assume that the true data-generating error distribution of ε_t is a non-Gaussian fat-tailed distribution. Specifically, we assume that ε_t follows the standard student- t distribution with degree of freedom $\nu = 5$, which is chosen be close to the estimated values in our empirical application in the next section. As in CCP, we consider the sample size $n = 500$ and 1000 replications. To examine the potential impact of misspecification, we assume that the misspecified model has the standard normal error distribution as in CCP but other parts of the model are correctly specified. We estimate both the correctly specified and the misspecified models from simulated data and study their differences.

3.2 | Results

We first examine the bias in the estimation of the parameters common to both the correctly and the incorrectly specified models, namely (σ_L, σ_H) , (α, τ) , and ρ . The results are presented in the upper panel of Tables 1 and 2 for scenarios $(\alpha, \tau) = (0.4, 0.5)$ and $(0.8, 0.7)$, respectively. First, for both scenarios, the biases for σ_L and σ_H are all positive across all levels of ρ . Importantly, while the biases under the correct specification are quite close to zero, those under the incorrect specification are substantially larger. This is expected because under the incorrect normal specification larger estimates of σ_L and σ_H are needed to compensate for the more frequent extreme observations generated by the fat-tailed Student- t error distributions. Second, for the first scenario, the biases for α under both specifications are negative, whereas for the second scenario they are positive under the correct specification but negative under the incorrect specification. What is important is that in absolute values the biases under the correct specification are always much smaller than those under the incorrect specification. Third, for the first scenario, the biases for τ under both specifications are negative. For the second scenario, they are all negative under the normal specification, whereas under the Student- t specification they are positive when ρ is close to 0 but gradually become negative when ρ approaches -1 . In absolute values, the correctly specified models again always have much smaller biases. Finally, the bias results for ρ are somewhat mixed between the two error specifications. For both scenarios, the biases for ρ under the correct specification are actually inferior to those under the incorrect specification when ρ is close to 0, but the opposite happens when ρ is close to -1 . While the biases for ρ tend to be small under the incorrect normal specification, they are quite clearly negative for some values of ρ under the correct student- t specification.

In addition to parameter estimation, testing the presence of endogenous regime changes and extracting the unobserved latent factor are two particularly useful inferential tasks in the framework of latent-factor-driven endogenous regime-switching models. We expect that misspecifications in the state-dependent dynamics may have negative impacts on the performance of the LR test considered by CCP on the hypothesis $\rho = 0$ and the accuracy of the extracted latent factor. We investigate the respective sizes and powers of the LR test under correct and incorrect specifications. The LR test statistic is given by

$$LR = 2(\ell(\hat{\theta}, \hat{\alpha}, \hat{\tau}, \hat{\rho}) - \ell(\tilde{\theta}, \tilde{\alpha}, \tilde{\tau}))$$

where ℓ denotes the log-likelihood function and the parameters with tildes and hats denote their ML estimates

TABLE 1 Simulation results
($\alpha = 0.4, \tau = 0.5$)

	ρ	0	-0.2	-0.4	-0.6	-0.8	-1
Bias for σ_H	Normal	0.0809	0.0769	0.0739	0.0642	0.0533	0.0405
	Student	0.0126	0.0100	0.0077	0.0042	0.0032	0.0040
Bias for σ_L	Normal	0.0208	0.0195	0.0177	0.0138	0.0092	0.0049
	Student	0.0023	0.0019	0.0025	0.0018	0.0016	0.0012
Bias for α	Normal	-0.6027	-0.5962	-0.5925	-0.5699	-0.4928	-0.4443
	Student	-0.4053	-0.3972	-0.3574	-0.3445	-0.3425	-0.3997
Bias for τ	Normal	-1.0150	-0.9705	-0.9139	-0.7821	-0.5943	-0.3925
	Student	-0.2582	-0.2147	-0.2689	-0.2570	-0.2356	-0.2366
Bias for ρ	Normal	0.0014	0.0518	0.0845	0.1024	0.0728	0.0259
	Student	-0.0085	-0.0639	-0.1124	-0.1007	-0.0660	0.0090
Power	Normal	0.0500	0.2420	0.7110	0.9510	0.9920	0.9980
	Student	0.0500	0.2500	0.8140	0.9910	1.0000	1.0000
AMSE	Normal	1.2787	1.3517	1.7995	2.4435	3.4531	4.4744
	Student	1.2756	1.2854	1.5403	1.8907	2.5443	3.3271

TABLE 2 Simulation results
($\alpha = 0.8, \tau = 0.7$)

	ρ	0	-0.2	-0.4	-0.6	-0.8	-1
Bias for σ_H	Normal	0.0683	0.0659	0.0641	0.0584	0.0507	0.0448
	Student	0.0034	0.0043	0.0034	0.0056	0.0062	0.0070
Bias for σ_L	Normal	0.0152	0.0140	0.0136	0.0122	0.0107	0.0086
	Student	0.0018	0.0017	0.0021	0.0028	0.0029	0.0026
Bias for α	Normal	-0.1837	-0.1583	-0.1452	-0.0936	-0.0435	-0.0358
	Student	0.0655	0.0671	0.0811	0.0778	0.0571	0.0018
Bias for τ	Normal	-0.8258	-0.7594	-0.7242	-0.6342	-0.5016	-0.4120
	Student	0.4131	0.3094	0.2053	-0.0328	-0.0921	-0.1339
Bias for ρ	Normal	0.0016	0.0148	0.0295	-0.0065	-0.0296	0.0298
	Student	0.0058	-0.1505	-0.2496	-0.2555	-0.1613	0.0042
Power	Normal	0.0500	0.1880	0.5590	0.8780	0.9690	0.9840
	Student	0.0500	0.2600	0.7310	0.9810	1.0000	1.0000
AMSE	Normal	2.4500	2.5818	3.2376	4.5288	6.9065	8.6650
	Student	2.4437	2.5401	2.9947	3.2362	3.5800	3.8081

with and without the no endogeneity restriction $\rho = 0$. Under the correct specification, $\theta = (\sigma_L, \sigma_H, v)$, and under the incorrect specification, $\theta = (\sigma_L, \sigma_H)$. We calculate the LR statistic from the estimated restricted (exogenous) and unrestricted (endogenous) models and repeat this for all the replications. Following CCP, we obtain the sizes of the test using the critical value from the $\chi^2(1)$ limiting distribution and the 5% size-adjusted powers using the empirical 95th percentile of the LRs from data generated under the null $\rho = 0$ as the critical value.

The empirical sizes of the LR test for the scenario $(\alpha, \tau) = (0.4, 0.5)$ are 0.081 and 0.062, respectively, under correct and incorrect specifications, and for the scenario $(\alpha, \tau) = (0.8, 0.7)$ the numbers are 0.084 and 0.086,

respectively. This suggests that the test slightly over rejects the null hypothesis and the misspecification considered in our setting does not appear to have substantial impact on the size of the test. We then present the size-adjusted power results from the two scenarios in middle panel of Tables 1 and 2, respectively. Generally speaking, the test is reasonably powerful under both specifications, with the powers increasing quite rapidly as ρ gets away from 0. However, as we expected, it is important to note that the size-adjusted power under the correct specification is always better than that under the incorrect specification, and not surprisingly their difference increases as the endogeneity level increases. Under the misspecification considered in our setting, the LR test has

the tendency to under-reject the null hypothesis of no endogeneity. This means that researchers may reach the conclusion of no endogenous regime changes more often than they should due to this type of misspecifications.² The implication of this should not be underestimated, because simulation results of CCP confirmed that ignored endogeneity can lead to quite substantial biases and deteriorated efficiency in the estimation of model parameters.

Finally, we examine the impact of misspecification on the quality of the extracted latent factor. An appealing advantage of the latent factor-driven regime-switching model is that once the model has been estimated, the researcher can extract the latent factor to study, for instance, its macroeconomic determinants and their joint dynamics to improve forecasting performance of their model. See Chang, Kwak, and Qiu (2019) for an example. Since the true sample path of (w_t) for each simulated time series is known, one natural way of assessing the accuracy of the extracted latent factor is to calculate the following Mean Square Error (MSE) for each simulated series and average them across all the replications. We let $i = 1, 2, \dots, 1000$ be the replication number for each simulation case and define MSE for each replication i as

$$MSE(i) = \frac{1}{n} \sum_{t=1}^n (w_{i,t} - \hat{w}_{i,t})^2$$

and the Average MSE (AMSE) across all replications as

$$AMSE = \frac{1}{1000} \sum_{i=1}^{1000} MSE(i)$$

Here, $w_{i,t}$ is true latent factor and $\hat{w}_{i,t}$ is the extracted value defined in (10) from an estimated model. We do this for both the correctly specified and the incorrectly specified models, and report the results in the bottom panel of Tables 1 and 2. We can clearly see that correctly specified models always produce smaller AMSE measures than incorrectly specified models. As the endogeneity level increases from $\rho = 0$ to -1 , the difference increases quite rapidly.

In summary, under our simulation setting, we find that the type of misspecification we consider, namely the ignored the fat-tailedness, generally leads to larger biases in the estimation of model parameters, reduced power of the LR test against endogeneity, and deteriorated accuracy of the extracted latent factor. All these may have some negative impacts on the application of the latent-factor-driven endogenous regime-switching framework to empirical economic and financial research, potentially rendering invalid conclusions. Our empirical application in the next session reveals some additional evidence.

4 | APPLICATION

4.1 | Data and models

To demonstrate further the benefit of being able to specify general state-dependent dynamics, we analyse the daily returns of the S&P 500 index based on essentially the same regime-switching volatility model studied in our simulation. We use the daily series of demeaned log returns (in percentage) of the S&P 500 index from January 3, 1928 to December 31, 2018 (22,856 observations). We further divide the full sample into an estimation sample from January 3, 1928 to December 29, 2006 (19,836 observations) and a forecasting sample from January 3, 2007 to December 31, 2018 (3,020 observations). We can see from Figure 1 that our data exhibits quite clear changing volatility levels through different time periods. Several periods of high volatilities can be clearly identified, including at the very least the Great Recession period from 1920s to 1930s, the financial crises of the late 1990s, and the more recent of crisis between 2007 and 2008, for example. As a relatively simple strategy for illustration purpose, we consider the same regime-switching volatility model (11) with alternative, namely the normal and the Student- t , error distributions.³ Slightly different from our simulation, we now also allow the degree of freedom parameter in the Student- t distribution to be regime-dependent, taking value ν_L or ν_H according to the volatility regime.

For comparison purposes, for both error specifications, we consider three model types in terms of regime-switching specification. The three types of models are the single-regime models, the exogenous regime-switching models ($\rho = 0$), and the endogenous regime-switching models ($\rho \neq 0$). Consequently, we have a total of six

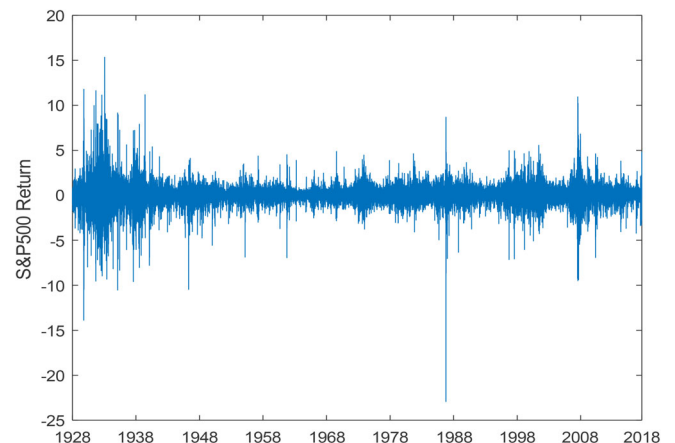


FIGURE 1 Daily S&P 500 index returns [Colour figure can be viewed at wileyonlinelibrary.com]

competing models to be investigated. All our models are estimated by ML from our estimation sample.

4.2 | In-sample results

We report our in-sample estimation results in Table 3. First of all, we observe that for models from the same switching type, the one with normal errors produces larger estimates of σ_H and σ_L , smaller estimates of α and τ , and a less negative estimate of ρ than the one with Student- t errors. This is consistent with the observation from our simulation, where the true data-generating processes have Student- t errors. Our initial investigation focuses on the evidence of endogenous regime-switching effect in the data. We begin with a comparison of the log-likelihood values across the three types of models. Our results show that for a given error distribution the endogenous regime-switching model always produces the highest log-likelihood, which is followed by the exogenous regime-

switching model, and the single-regime model always has the lowest log-likelihood. Traditional test statistics cannot be used for testing whether there is one regime or two regimes, because the parameters related to the second regime of the process are not identified under the null hypothesis of no regime-switching, a feature known as the Davies' Problem (c.f. Davies, 1977, 1987). Nevertheless, studies such as Hamilton and Susmel (1994), Gray (1996), and Choi (2009) argued that at least informally the standard LR statistic can give some indication of the relative performance of competing models in terms of their goodness-of-fit to the data. From the log-likelihood value reported in Table 3, we can calculate the LRs between the exogenous regime-switching model and the single-regime model to be 9764.24 and 3203.84 for the student- t and the normal error specifications, respectively. Arguably the introduction of two regimes enormously improves the goodness-of-fit to the data.

To formally test the statistical significance of the observed likelihood improvements, we resort to parametric

TABLE 3 Estimation results for daily S&P 500 index returns

Parameter	Normal			Student		
	Single	Exogenous	Endogenous	Single	Exogenous	Endogenous
σ_H		2.2585 (0.0308)	2.264 (0.0039)		1.3488 (0.0792)	1.3757 (0.3280)
σ_L	1.1806 (0.0055)	0.7035 (0.0048)	0.7032 (0.0182)	0.6398 (0.0023)	0.5736 (0.0446)	0.5768 (0.0570)
ν_L				2.5485 (0.0014)	6.1128 (0.2557)	6.1872 (0.4542)
ν_H					3.6768 (0.0966)	3.7185 (0.5910)
α		0.9969 (0.0004)	0.9960 (0.0009)		0.9997 (0.0004)	0.9996 (0.0003)
τ		11.0007 (0.6788)	9.8100 (0.1234)		23.7776 (1.5464)	19.7607 (2.2122)
ρ			-0.7216 (0.0025)			-0.9828 (0.0350)
p_{LL}		0.9880	time-varying		0.9953	time-varying
p_{HH}		0.9492	time-varying		0.9871	time-varying
log-likelihood	-31,438.60	-26,556.48	-26,526.62	-27,623.76	-26,021.84	-25,998.72
LR for single regime	9,764.24			3,203.84		
bootstrap p -value	0.000			0.000		
LR for $\rho = 0$		59.72			46.24	
bootstrap p -value		0.000			0.000	
LR for distribution	7,629.68	1,069.28	1,055.80			
bootstrap p -value	0.000	0.000	0.000			

bootstrapping. We simulate 1,000 replications of artificial data from the estimated single-regime model. For each replication, we estimate models under the null and the alternative by ML and calculate the LR. From the empirical distribution of the 1,000 bootstrap LRs, we can find the bootstrap p -value for the observed LR from the original sample. Unlike asymptotic tests such as Davies (1987), Hansen (1992, 1996), Garcia (1998), the parametric bootstrap procedure is easy to implement and relies on the finite sample empirical distribution rather than large sample approximation. We report the resulting bootstrap p -values underneath the sample LR statistics. As expected, our bootstrap p -values for the above two LRs are both zero, meaning that when the true data-generating model is the single-regime model, none of the simulated LRs exceeds the LR from the original sample. Clearly, this is very strong evidence against the single-regime model.

Comparing the two regime-switching models, we find that the estimated endogeneity parameters are -0.7216 and -0.9828 for the two error specifications, respectively, revealing quite strong levels of endogenous regime changes. The two LR statistics between the endogenous and exogenous models are 59.72 and 46.24, respectively, leading to strong rejections of no endogeneity according to the limiting distribution. To further test the significance of endogenous regime-switching effect in our data, we again resort to parametric bootstrapping. The resulting two bootstrap p -values are both zero again, confirming statistical significance of the endogeneity. This is consistent with the empirical findings in CCP for stock returns and many existing works, including for example Choi (2009) on short term interest rates, Bu et al. (2017) on stock market volatilities.

We now compare our estimation results from different error specifications. First, we observe that for models of the same switching type, the one with the Student- t error distribution always produces a substantially higher log-likelihood value than the one with the normal error distribution. This is an expected in-sample result because the Student- t distribution is more flexible and becomes equivalent to the normal distribution when the degree of freedom tends to infinity. Unsurprisingly, the LRs between models with alternative error specifications are as high as 7629.68, 1069.28, and 1055.80, respectively, for the three switching types. More formally, the respective p -values obtained from parametric bootstrap tests on these LRs are all zero, confirming the significance of the differences between the two error specifications. Combining the results from our tests on regime-switching effects and those on error specifications, we can conclude quite confidently that among the six models under our consideration, the endogenous regime-switching model with the Student- t error distribution is by far the best fitting model for our data.

Endogenous regime-switching models have time-varying state transition probabilities that are functions of the lags of the time series variable y_t . For the volatility model in (11), the state transition probabilities are functions of y_{t-1} only and hence we can denote them as $P_{LL}(y_{t-1})$ and $P_{HH}(y_{t-1})$. Figure 2 plots the estimated values of $P_{LL}(y_{t-1})$ and $P_{HH}(y_{t-1})$ across the whole estimation sample period for the two endogenous regime-switching models together with the estimated constant transition probabilities P_{LL} and P_{HH} for the two exogenous regime-switching models. We can see that, consistent with the results in CCP, the time-varying probabilities are substantially different from the constant ones almost at all times, suggesting that ignoring endogeneity in regime changes can lead to considerable bias in the estimation of state transition probabilities. Recall that the endogenous regime-switching effect is significant regardless of the error specification according to our LR tests above. Basically, this means that the time variation in the state transition probabilities presented in Figure 2 is statistically significant regardless of the error specification.

Meanwhile, we can also observe that the time-varying state transition probabilities $P_{LL}(y_{t-1})$ and $P_{HH}(y_{t-1})$ for models with normal errors are substantially more variable and often much lower than those for models with Student- t errors. To understand this observation properly, we plot the estimated $P_{LL}(y_{t-1})$ and $P_{HH}(y_{t-1})$ as a function of y_{t-1} across its empirical support for models with different errors in Figure 3. We can see that for $y_{t-1} > 0$, $P_{LL}(y_{t-1})$ is very close to unity for both specifications. When the process is in state L at time $t-1$, a positive shock $u_{t-1} = y_{t-1}/\sigma_L$ or $\tilde{u}_{t-1} = y_{t-1}/\sigma_L$ combined with a negative ρ leads to a negative w_t , which in turns leads to a high probability of staying in state L . In contrast, a negative shock leads to a much lower probability of staying in state L . The opposite situation applies to $P_{HH}(y_{t-1})$. Meanwhile, models with normal errors have lower $P_{LL}(y_{t-1})$ and $P_{HH}(y_{t-1})$ than those with Student- t errors. This means that the model with normal errors produced more frequent regime changes than did the model with Student- t errors. A possible explanation is that data from potentially fat-tailed distributions have more extreme and more variable observations than what the normal distribution can predict. Consequently, misspecified models with Gaussian errors tend to over-estimate the probabilities of regime changes to compensate this additional variation and extreme observations.

4.3 | Out-of-sample results

We now examine the out-of-sample performance, focusing on the two endogenous regime-switching models. As

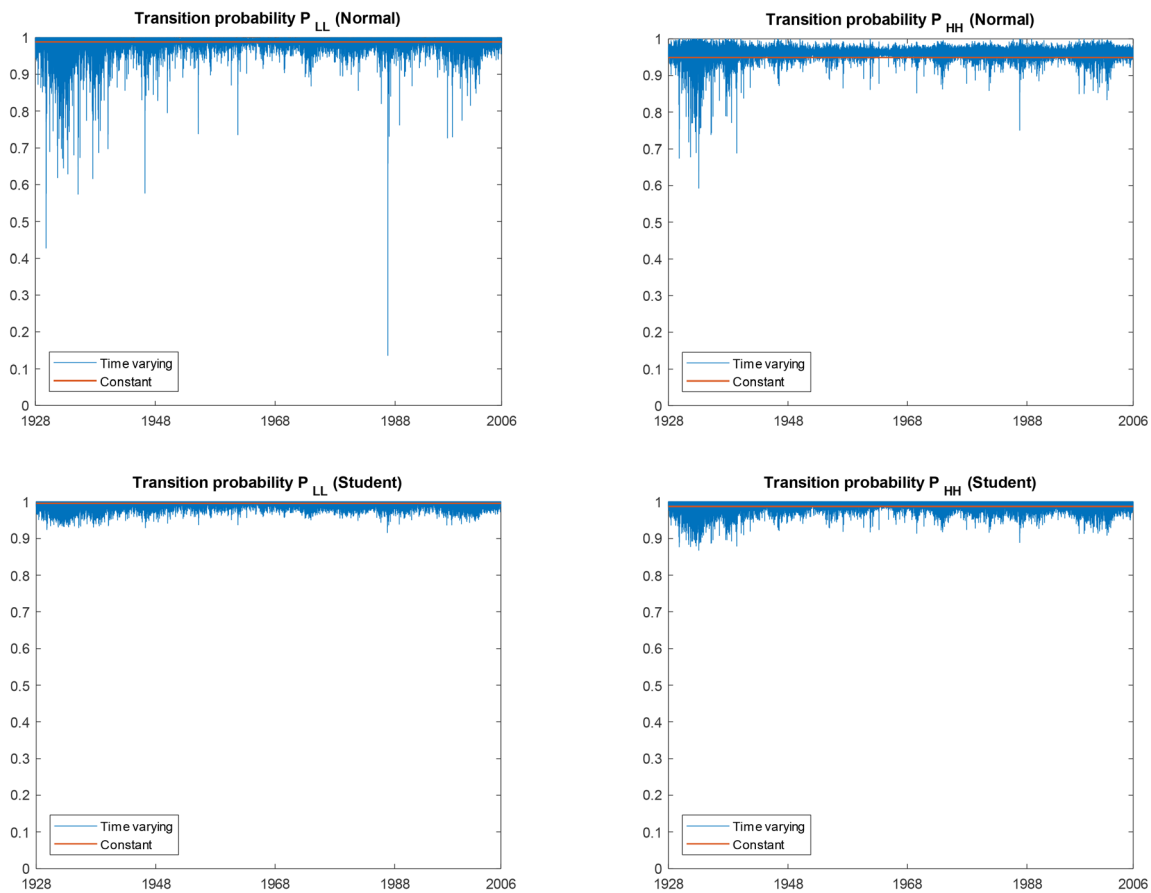


FIGURE 2 Time series of estimated state transition probabilities [Colour figure can be viewed at wileyonlinelibrary.com]

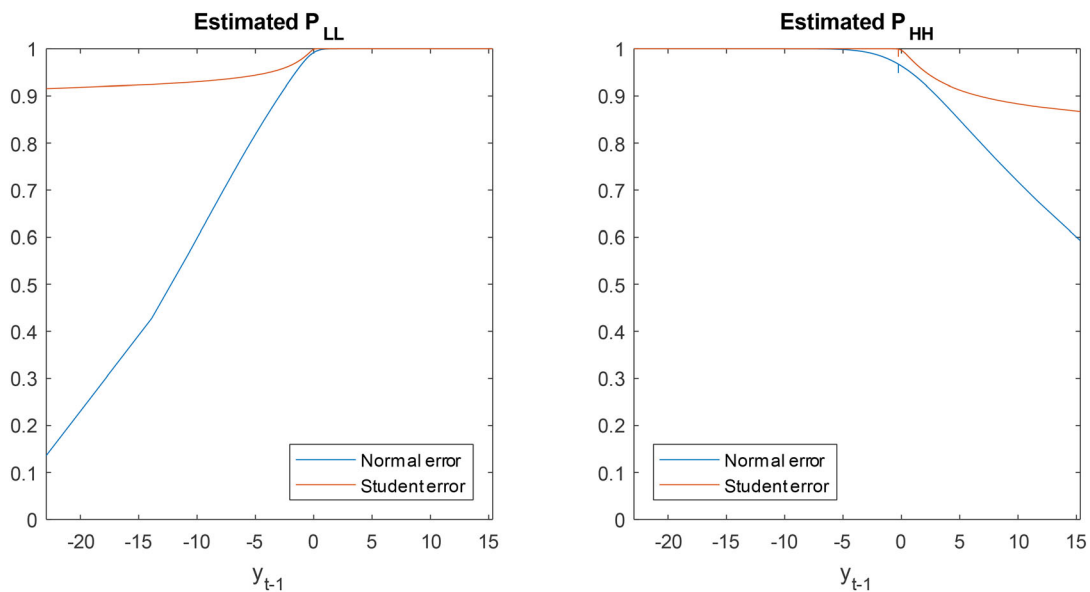


FIGURE 3 Estimated state transition probabilities as a function of y_{t-1} [Colour figure can be viewed at wileyonlinelibrary.com]

a standard way of evaluating volatility forecasting performance, we use the squared return as the proxy of the actual volatility. Various forecasting criteria or loss functions may be considered to assess the predictive accuracy

of volatility models, and the value of loss function may be affected by the choice of the proxy of actual volatility heavily. Patton (2011) shows that two popular loss functions, that is, MSPE and Gaussian Quasi-Likelihood

(QLIKE), are more robust to the imperfect volatility proxies. These two loss functions are given by

$$MSPE = \frac{1}{N} \sum_{i=1}^N (\sigma_i^2 - \hat{\sigma}_i^2)^2,$$

and

$$QLIKE = \frac{1}{N} \sum_{i=1}^N \left(\log \hat{\sigma}_i^2 + \frac{\sigma_i^2}{\hat{\sigma}_i^2} \right)^2$$

where σ_i^2 and $\hat{\sigma}_i^2$ are the proxy actual value and the one-step-ahead rolling sample volatility forecast, respectively, and N is the total number of out-of-sample volatility forecasts. We find that the MSPE and QLIKE from our endogenous regime-switching model with Student- t errors are 27.89 and 0.96, respectively, compared with 28.16 and 0.99 from the model with normal errors. To formally examine the significance of such differences, we consider a one-sided DM test of Diebold and Mariano (1995), using the endogenous regime-switching model with normal errors as the benchmark, on the null hypothesis that the two models produce equal forecasts. The results are statistically significant, with p -values being 0.0471 and 0.0436, indicating statistical significance at 5% significance level. Based on this evidence, our endogenous regime-switching model with Student- t errors has better out-of-sample performance as far as our forecasting criteria are concerned. Combined with our in-sample results, we may reach the important conclusion that the ability to model regime-switching with endogenous regime changes and non-Gaussian state-dependent conditional distributions can be extremely useful in empirical econometric modelling.

Finally, we demonstrate the importance and potential economic implications of making appropriate distributional assumptions on the state-dependent dynamics in endogenous regime-switching models. We investigate this from the perspective of financial risk management. Since our focus is on the fat-tailedness in financial returns, we compare our models in terms of a suitable tail risk measure. To exploit as much information in the tails of our forecasting distributions as possible, we consider the CTE, defined as the expected value of the loss given that the loss falls in the upper tail of the distribution,⁴ as our coherent tail risk measure. Artzner, Delbaen, Eber, and Heath (1999) and others argued that CTE gives better results than does the quantile measure when comparing risks, because it utilizes the whole tail of the distribution beyond the quantile, rather than the single quantile point. We assume for simplicity that the S&P 500 index is

the single investment asset under our consideration, and since our models are symmetric, for comparison purpose it is sufficient to consider CTE based on one of the tails. As such, for a given confidence level $1 - \alpha$ where $\alpha \in (0, 1)$, we define our CTE in terms of the daily S&P 500 return X as

$$CTE(\alpha) = E[X|X < F^{-1}(\alpha)] = \frac{1}{\alpha} \int_{-\infty}^{F^{-1}(\alpha)} xf(x)dx$$

where $f(x)$ is the forecast of the probability density function of the S&P 500 return and $F(x)$ is the corresponding cumulative distribution function.

To contrast the difference in forecasting CTE under alternative error distributional assumptions, that is, the normal distribution and the Student- t distribution, and focus on the deep tails, we choose the 99% confidence level or $\alpha = 0.01$. This amounts to comparing the performance of alternative model specifications in forecasting the top or bottom 1% probability mass in the tails of the S&P 500 return distribution. We plot in Figure 4 the one-period-ahead rolling sample CTE forecasts in the left tail produced by our endogenous regime-switching model with alternative error specifications as well as the actual return series in our forecasting sample period.

First, we can clearly see that not surprisingly our CTE forecasts vary quite considerably over time, closely in line with the variation of the return itself. Second, the CTE forecasts produced by the model with Student- t errors are almost always more extreme than those produced by the model with normal errors. Their differences are more substantial in high volatility regimes than in low

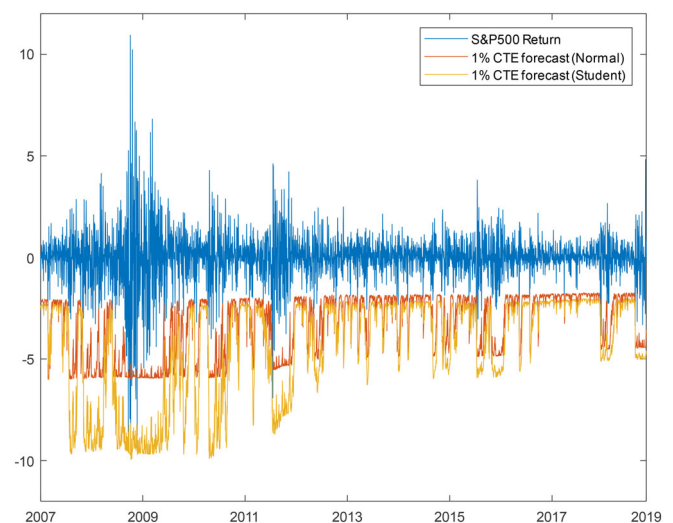


FIGURE 4 S&P 500 returns and CTE forecasts [Colour figure can be viewed at wileyonlinelibrary.com]

volatility regimes. This means that models with the normal error specification with no fat tails tend to under-estimate the tail risks most of the time relative to models with the more adequate Student-*t* error specification with fat tails. This under-estimation can be considerable particularly during crisis periods where extreme movements of the market are more frequently. To formally confirm the significance of the difference in the CTE forecasts produced by alternative models, we resort again to the one-sided DM test of Diebold and Mariano (1995), using the model with normal errors as our benchmark, on the null hypothesis that the two models produce equal CTE forecasts. The resulting statistic is as large as 38.62 and the corresponding *p*-value is negligibly small. Hence, our test strongly rejects the null, which means that our fat-tailed Student-*t* error specification produced significantly more extreme CTE forecasts than does the normal error specification for the S&P 500 returns data. The important implication from this example is that in the framework the endogenous regime-switching models, the Gaussian assumption may significantly under-estimate economic or financial risks compared to alternative more general assumptions.

5 | CONCLUSION

We extend the class of latent factor-driven endogenous regime-switching models of Chang et al. (2017) by allowing the state-dependent dynamics to be general. Our generalization makes this class of models applicable to a much wider range of realistic situations, where non-Gaussian state-dependent transitions are considered to be essential.

Focusing on the potential impacts of fat-tailed distributions, which is a prominent feature of most financial return series, on the estimation and inference of this class of models, we conducted a simulation study to examine the consequences of misspecification. We found that in addition to the well-anticipated larger biases in the estimation of almost all model parameters, misspecified state-dependent dynamics also lead to power losses in the LR test against endogenous regime switches, uniformly across different levels of true endogeneity, which effectively increases the chance of not detecting crucial time-variations in the transition probabilities of the estimated models. Moreover, we found misspecified state-dependent dynamics may also deteriorate the accuracy of the extracted unobserved latent factors, which in turn could undermine the quality and robustness of any subsequent analyses based on it.

To illustrate our proposed approach empirically and demonstrate further the benefit of allowing the state-

dependent dynamics to be more flexible, we analyse the daily returns of the S&P 500 equity index based on a regime-switching volatility model. We found quite strong evidence of endogenous regime changes and that models with the fat-tailed student-*t* errors outperformed those with Gaussian errors in terms of both in-sample and out-of-sample performances. Our non-Gaussian model also produced significantly more extreme tail risk forecasts, which can have quite important risk management implications.

ENDNOTES

- ¹ We conducted simulation for different values of ρ between 0 and -1 on a grid of size 0.1 but chose to report results on a grid of size 0.2 in Tables 1 and 2 below to save space.
- ² Clearly, this is true only as far as our simulation design is concerned. We expect that different types or degrees of misspecifications may have different levels of impact on the size and power of the test, which we leave for future research.
- ³ To keep our analysis relatively focused, we do not account for potentially empirical features such as volatility clustering and asymmetric error distribution. We leave them for future investigation.
- ⁴ CTE is also known as Tail Conditional Expectation (TCE), Conditional Value-at-Risk (CVaR), Tail Value-at-Risk (TVaR), and Expected Shortfall (ES).

REFERENCES

- Artzner, P., Delbaen, F., Eber, J. M., & Heath, D. (1999). Coherent measures of risk. *Mathematical Finance*, 9(3), 203–228.
- Bollen, N. P., Gray, S. F., & Whaley, R. E. (2000). Regime-switching in foreign exchange rates: Evidence from currency option prices. *Journal of Econometrics*, 94, 239–276.
- Bu, R., Cheng, J., & Hadri, K. (2017). Specification analysis in regime-switching Continuous-time diffusion models for market volatility. *Studies in Nonlinear Dynamics and Econometrics*, 21(1), 65–80.
- Chang, Y., Choi, Y., & Park, J. Y. (2017). A new approach to model regime switching. *Journal of Econometrics*, 196(1), 127–143.
- Chang, Y., Kwak, B., and Qiu, S. (2019) U.S. Monetary-Fiscal Regime Changes in the Presence of Endogenous Feedback in Ploicy Rules. *Working Paper*.
- Choi, S. (2009). Regime-switching univariate diffusion models of the short-term interest rate. *Studies in Nonlinear Dynamics and Econometrics*, 13(1), 4.
- Davies, R. B. (1977). Hypothesis testing when a nuisance parameter is present only under the alternative. *Biometrika*, 64, 247–254.
- Davies, R. B. (1987). Hypothesis testing when a nuisance parameter is present only under the alternatives. *Biometrika*, 74(1), 33–43.
- Diebold, F., Lee, J., & Weinbach, G. (1994). Regime switching with time-varying transition probabilities. In C. Hargreaves (Ed.), *Nonstationary Time Series Analysis and Cointegration* (pp. 283–302). Oxford: Oxford University Press.
- Diebold, F. X., & Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business and Economic Statistics*, 13(3), 253–263.

- Garcia, R. (1998). Asymptotic null distribution of the likelihood ratio test in Markov switching models. *International Economic Review*, 39(3), 763–788.
- Garcia, R., & Perron, P. (1996). An analysis of the real interest rate under regime shifts. *Review of Economics and Statistics*, 78(1), 111–125.
- Goutte, S., & Zou, B. (2013). Continuous time regime switching model applied to foreign exchange rate. *Mathematical Finance Letters*, 2013, 8.
- Gray, S. F. (1996). Modeling the conditional distribution of interest rates as a regime-switching process. *Journal of Financial Economics*, 42, 27–62.
- Hamilton, J. (1988). Rational-expectations econometric analysis of changes in regime: An investigation of the term structure of interest rates. *Journal of Economic Dynamics and Control*, 12, 385–423.
- Hamilton, J. (1989). A new approach to economic analysis of non-stationary time series. *Econometrica*, 57, 357–384.
- Hamilton, J., & Susmel, R. (1994). Autoregressive Conditional heteroskedasticity and changes in regime. *Journal of Econometrics*, 64, 307–333.
- Hansen, B. E. (1992). The likelihood ratio test under nonstandard conditions: Testing the Markov switching model of GNP. *Journal of Applied Econometrics*, 7, S61–S82.
- Hansen, B. E. (1994). autoregressive conditional density estimation. *International Economic Review*, 35, 705–730.
- Hansen, B. E. (1996). Erratum: The likelihood ratio test under non-standard conditions: Testing the Markov switching model of GNP. *Journal of Applied Econometrics*, 11(2), 195–198.
- Jondeau, E., & Rockinger, M. (2003). Conditional volatility, skewness, and kurtosis: existence, persistence, and comovements. *Journal of Economic Dynamics and Control*, 27(10), 1699–1737.
- Kasahara, H., & Shimotsu, K. (2019). Asymptotic properties of the maximum likelihood estimator in regime switching econometric models. *Journal of Econometrics*, 208, 442–467.
- Kim, C. J. (1994). Dynamic linear models with markov-switching. *Journal of Econometrics*, 60, 1–22.
- Kim, C. J., Piger, J., & Startz, R. (2008). Estimation of Markov regime-switching regression models with endogenous switching. *Journal of Econometrics*, 143, 263–273.
- Patton, A. J. (2011). Volatility forecast comparison using imperfect volatility proxies. *Journal of Econometrics*, 160, 246–256.

How to cite this article: Bu R, Cheng J, Jawadi F. A latent-factor-driven endogenous regime-switching non-Gaussian model: Evidence from simulation and application. *Int J Fin Econ*. 2020;1–16. <https://doi.org/10.1002/ijfe.2192>