Specification Analysis in Regime-Switching Continuous-Time Diffusion Models for Market Volatility

Ruijun Bu*
University of Liverpool

Jie Cheng
Xi'an Jiaotong-Liverpool University

Kaddour Hadri
Queen's University Belfast

*Corresponding author: Department of Economics, Finance and Accounting, Management School, Chatham Street, Liverpool, L69 7ZH, UK, Tel: +44-151-795-3122, Fax: +44-151-795-3004, Email: ruijunbu@liv.ac.uk.

Specification Analysis in Regime-Switching Continuous-Time Diffusion Models for Market Volatility

Abstract

We examine model specification in regime-switching continuous-time diffusions for modelling S&P 500 Volatility Index (VIX). Our investigation is carried out under two nonlinear diffusion frameworks, the NLDCEV and the CIRCEV frameworks, and our focus is on the nonlinearity in regime-dependent drift and diffusion terms, the switching components, and the endogeneity in regime changes. While we find strong evidence of regime-switching effects, models with a switching diffusion term capture the VIX dynamics considerably better than models with only a switching drift, confirming the presence and importance of volatility regimes. Strong evidence of nonlinear endogeneity in regime changes is also detected. Meanwhile, we find significant nonlinearity in the regime-dependent diffusion specification, suggesting that the nonlinearity in the VIX dynamics cannot be accounted for by regime-switching effects alone. Finally, we find that models based on the CIRCEV specification are significantly closer to the true data generating process of VIX than models based on the NLDCEV specification uniformly across all regime-switching specifications.

JEL Classification: C22, C24, C52, C58

Keywords: Volatility Index, Regime-Switching Model, Nonlinear Diffusion, Constant Elasticity Volatility, Endogeneity, Maximum Likelihood Estimation.

1. Introduction

The Volatility Index (VIX) was introduced by the Chicago Board Options Exchange (CBOE) in 1993 to measure the market's expectation of volatility based on the S&P 100 Index option prices, and it has since been an important indicator of implied short-term U.S. market volatility. In 2003, options on the S&P 500 Index were included in the calculation of VIX covering a wide range of strike prices. In 2004, VIX futures and options were also introduced, making volatility a tradable asset. Due to its nature of measuring market volatility, the VIX is termed as the "investor fear gauge" by Whaley (2000). Proper understanding of the VIX dynamics is key to the success of many financial and economic activities that rely on it.

VIX generally exhibits mean-reversion despite its volatile upward movements during some periods after the 2008 financial crisis. It is not clear, however, whether VIX reverts to the same constant level over time or whether its evolution is time-homogeneous. Guo and Wohar (2006) investigated the stability of the mean of VIX by performing multiple structural breaks test of Bai and Perron (1998) and identified multiple regimes. Chen and Huang (2014) examined VIX Exchange Traded Funds data and also found evidence structural breaks. Their findings, amongst those of others, are indicative of possible regime-switching features of VIX.

There is a large literature supporting the existence of stochastic regime changes in the evolution of financial and economic variables. Choi (2009) proposed a regime-switching univariate diffusion model to describe the dynamics of the U.S. short-term interest rates and found strong evidence of high and low volatility regimes and time-varying regime-switching probabilities. Goutte and Zou (2013) employed a modified Cox-Ingersoll-Ross (CIR) model to study exchange rates and documented that their regime-switching models match the reality much better than their single-regime model. Chevallier and Goutte (2014) considered a regime-switching jump diffusion process and found evidence of regime changes in the daily returns of seven international stock markets. Most recently, Bu et al. (2016) showed that regime-switching transformed diffusion models are effective in capturing additional variability in the dynamics of short-term interest rates.

While there have been numerous studies on regime-switching models, very few papers introduced regime shifts into continuous-time diffusion models to study the VIX dynamics. Among them, Amengual and Xiu (2012) proposed a joint model for the S&P 500 index and the VIX where the central tendency parameter for the spot volatility is allowed to be regime-switching. Papanicolaou and Sircar (2014) proposed a regime-switching Heston model for pricing VIX derivatives[1]. These studies almost exclusively considered regime-switching continuous-time affine (linear) processes. While affine models are extremely popular for pricing contingent claims due to their analytic tractability, the validity of the affine assumption is usually taken for granted.

[1] Mencia and Sentana (2013) and Song and Xiu (2014) also studied volatility models with stochastic components, but their models allowed these components to change continuously.

This paper contributes to the literature modelling the VIX dynamics by studying model specification in regime-switching continuous-time diffusion models. Our focus is on the nonlinearity in regime-dependent drift and diffusion terms, the switching components, and the potential nonlinear endogeneity in regime changes. *First*, our regime-switching continuous-time diffusion models are two-factor models. One is the continuously evolving regime-dependent diffusion process and the other is the regime variable changing discretely. Our models are economically intuitive, flexible, and easy to implement. Moreover, we consider two general nonlinear diffusions as our regime-dependent processes. The first is the Nonlinear Drift Constant Elasticity Volatility (NLDCEV) process introduced by Aït-Sahalia (1996b), and the second is the CIR-reducible Constant Elasticity Volatility (CIRCEV) process proposed by Bu et al. (2011). The two nonlinear diffusions are known to be able to generate desirable nonlinearity in both the drift and diffusion terms and encompass linear models (e.g. the CIR process) as special cases. Therefore, our frameworks allow us to test the validity of the nested linear (affine) models. *Second*, we test a number of restricted specifications in terms of whether the drift, the diffusion, or both terms are allowed to be regime-dependent. These tests shed light on whether stochastic central tendency, stochastic volatility or both are important for driving the VIX dynamics. *Third*, studies including Choi (2009), Chang et al. (2014), Bu et al. (2016) all reported evidence of endogeneity in regime changes. We investigate whether regime changes in the VIX dynamics, if any, are endogenously driven. We propose a flexible endogenous specification to pick up potential nonlinear endogenous dependence between regime-switching probabilities and the level of VIX. *Fourth*, the comparison between our two competing nonlinear diffusion specifications is a non-trivial issue due to their overlapping and dynamic nature. Using a combination of informal and formal criteria, we examine their relative closeness to the true data generating process (DGP) in our regime-switching context. *Finally,* our study does not suffer from discretization bias[2]. The CIRCEV specification has closed-form exact transition density by construction. For the NLDCEV specification, we use the highly accurate closed-form expansion method by Aït-Sahalia (2002) to approximate the unknown true transition density. The likelihood function of our regime-switching models are obtained by applying the recursive algorithm developed by Hamilton (1989).

To achieve our goals, we consider regime-switching models based on the CIR, the NLDCEV and the CIRCEV specifications. The linear CIR specification is included since it is a common special case of the two nonlinear frameworks. Hence, the linear restriction implied by the CIR specification can be tested under both nonlinear frameworks. In addition, for each of the three diffusion specifications, a total of seven distinct regime-switching specifications are considered for estimation. These purposefully specified models form a useful basis for examining the aforementioned specification issues in regime-switching diffusion models for VIX.

[2]See Hsiao and Semmler (2011) for a detailed discussion on the effects of discretization on continuous-time models.

Daily series of VIX data from January 2, 1990 to March 20, 2015 are used for estimating our models. Our main empirical findings can be summarized as follows: *First*, across all three diffusion specifications, we find strong evidence supporting the presence of two regimes in the VIX dynamics based on enormously improved likelihoods and information criteria. *Second*, all exogenous regime-switching models are strongly rejected against corresponding endogenous regime-switching models with a flexible time-varying transition probabilities. *Third*, drift-switching only models and diffusion-switching only models are both rejected against more general both-switching models. Nevertheless, diffusion-switching only models captures the VIX dynamics considerably better than drift-switching only models. *Forth*, across all seven regime-switching specifications, the linear CIR specification is strongly rejected under both nonlinear frameworks. *Finally*, using a combination of informal and formal comparisons, we find that models based on the more tractable CIRCEV specification outperform models based on the NLDCEV specification uniformly across all regime-switching specifications. The last point has particular implications for practical users interested in nonlinear diffusions, since our results suggest that the CIRCEV process which has a closed-form transition density can be a powerful alternative to the less tractable NLDCEV process whose transition density must be approximated.

The rest of this paper is organized as follows. In Section 2, we briefly introduce the two general continuous-time nonlinear diffusion frameworks and explain how to evaluate their transition densities. In Section 3, we outline the details of our regime-switching mechanism and our endogenous regime-switching specification. Empirical results based on the VIX data are presented in Section 4. Section 5 concludes.

2. Continuous-Time Nonlinear Diffusion Models

The basic dynamics for a univariate continuous-time parametric diffusion process $\{Y_t, t \geq 0\}$ is typically described by the following Stochastic Differential Equation (SDE)

$$dY_t = \mu_Y (Y_t; \psi) \, dt + \sigma_Y (Y_t; \psi) \, dW_t \tag{1}$$

where $\mu_Y (y; \psi)$ and $\sigma_Y^2 (y; \psi)$ are the instantaneous drift and diffusion terms, and $\{W_t, t \geq 0\}$ is a standard Brownian motion. Since the dynamics of Y_t is completely determined by $\mu_Y (y; \psi)$ and $\sigma_Y^2 (y; \psi)$, the main focus of diffusion modelling is on the specification these two terms and the estimation of ψ. Well known examples in finance include Merton (1973), Black and Scholes (1973), Vasicek (1977), Cox et al. (1985), Duffie and Kan (1996), Aït-Sahalia (1996b), Conley et al. (1997), Ahn and Gao (1999), and Bu et al. (2011)[3].

The diffusion term $\sigma_Y^2 (y; \psi)$ determines the conditional variance of the instantaneous change dY_t. When $\sigma_Y^2 (y; \psi) = \sigma^2 y^{2\theta}$, the process is said to have Constant

[3]Nonparametric and semiparametric diffusions have also been proposed in the literature. Notable examples include Aït-Sahalia (1996a), Stanton (1997), Jiang and Knight (1997), Bandi and Phillips (2003), Kristensen (2010), and most recently Bu et al. (2015).

Elasticity Volatility (CEV) where θ is the elasticity of the instantaneous volatility. The CEV specification was first introduced by Chan et al. (1992) who considered a linear drift in their study on the U.S. short-term interest rates. It was further studied by Aït-Sahalia (1996b) and Conley et al. (1997) and both promoted the use of a nonlinear drift to improve the mean reversion effect. The CEV specification is extremely parsimonious and the two parameters can usually be estimated very accurately. Empirical evidence also suggests that the CEV specification is capable of fitting the conditional instantaneous volatility adequately. Therefore, it has been considered by many empirical studies, including Gallant and Tauchen (1998), Durham (2003), Choi (2009), Bu et al. (2011, 2016).

2.1. Nonlinear Drift CEV (NLDCEV) Model

The NLDCEV diffusion process considered in this paper is specified as

$$dY_t = (\alpha_{-1}Y_t^{-1} + \alpha_0 + \alpha_1 Y_t + \alpha_2 Y_t^2)dt + \sigma Y_t^{\theta} dW_t \tag{2}$$

which is a generalization of the linear drift CEV model (also known as the CKLS model) by Chan et al. (1992). For the CKLS model, the strength of mean reversion is the same for all levels of Y_t. However, several empirical studies (e.g. Aït-Sahalia 1996b) concluded that the linear drift fails to generate strong enough pull at low or high levels of Y_t and hence suggested the above nonlinear drift model for a wider range of applications. When $\alpha_{-1} > 0$ and $\alpha_2 < 0$, the drift specification generates stronger mean reversion. The speed of Y_t returning to its long-run average value is given by $-\alpha_{-1}/Y_t^2 + \alpha_1 + 2\alpha_2 Y_t$. Conley et al. (1997) provided a detailed analysis of the dynamic properties of this process and derived specific parameter restrictions for the stationarity of the model. The NLDCEV model is general enough to nest several existing models such as the Vasicek (1977) model ($\alpha_{-1} = \alpha_2 = \theta = 0$), the CIR model ($\alpha_{-1} = \alpha_2 = 0$ and $\theta = 1/2$), and the CKLS model ($\alpha_{-1} = \alpha_2 = 0$).

Maximum Likelihood (ML) is usually the preferred method of estimation because of its efficiency gain. However, the transition density of a continuous-time diffusion is usually unavailable in closed-form except for a few rare cases (e.g. Black and Scholes 1973, Vasicek 1977, and Cox et al. 1985). In order to use ML, the transition density of the NLDCEV model must be approximated. Among many approximation techniques (c.f. Durham and Gallant 2002), the closed-form expansion method by Aït-Sahalia (2002) is widely accepted as one of the most efficient and accurate density approximation methods.

This method first employs the so-called Lamperti transform

$$X \equiv \gamma(Y;\psi) = \int^{Y} \frac{1}{\sigma_Y(\omega;\psi)} d\omega$$

to transform the original diffusion Y_t to a unit diffusion X_t, i.e.

$$dX_t = \mu_X(X_t;\psi)dt + dW_t$$

where

$$\mu_X(x;\psi) = \frac{\mu_Y(\gamma^{-1}(x;\psi);\psi)}{\sigma_Y(\gamma^{-1}(x;\psi);\psi)} - \frac{1}{2}\frac{\partial \sigma_Y}{\partial y}(\gamma^{-1}(x;\psi);\psi)$$

The transition density of X_t can then be approximated more reasonably by the Hermite expansion. Aït-Sahalia (2002) showed that the Kth order approximation of $p_X(\Delta, x|x_0;\psi)$ can be written as

$$\begin{aligned} & p_X^{(K)}(\Delta, x|x_0;\psi) \\ = \;& \Delta^{-\frac{1}{2}}\phi\left(\frac{x-x_0}{\Delta}\right)\exp\left(\int_{x_0}^{x}\mu_X(\omega;\psi)d\omega\right)\sum_{k=0}^{K} c_X^{(k)}(x|x_0;\psi)\frac{\Delta^k}{k!} \end{aligned} \tag{3}$$

where $\phi(\cdot)$ is the standard normal density function and $c_X^{(k)}(x|x_0;\psi) = 1$. The other coefficients are determined recursively by

$$\begin{aligned} & c_X^{(k)}(x|x_0;\psi) \\ = \;& k(x-x_0)^{-k}\int_{x_0}^{x}(\omega-x_0)^{k-1}\left\{\lambda_X(\omega;\psi)c_X^{(k-1)}(\omega|x_0;\psi) + \frac{1}{2}\frac{\partial^2 c_X^{(k-1)}(\omega|x_0;\psi)}{\partial\omega^2}\right\}d\omega \end{aligned}$$

with $\lambda_X(x;\psi) = \left[\mu_X(x;\psi)^2 + \partial\mu_X(x;\psi)/\partial x\right]/2$. Consequently, given $p_X^{(K)}(\Delta, x|x_0;\psi)$ in (3), we can obtain the Kth order approximation of the required transition density $p_Y^{(K)}(\Delta, y|y_0;\psi)$ by changing variable from X to Y as

$$p_Y^{(K)}(\Delta, y|y_0;\psi) = \frac{p_X^{(K)}(\Delta, \gamma(y;\psi)|\gamma(y_0;\psi);\psi)}{\sigma_Y(y;\psi)}$$

In practice, $K = 1$ gives very precise density function with the usual values of Δ (e.g. daily, weekly or monthly) (c.f. Aït-Sahalia 1999). In this paper, we use a daily data set (i.e. $\Delta = 1/252$) and hence use the approximate density function with $K = 1$.

2.2. CIR-Reducible CEV (CIRCEV) Model

The NLDCEV specification is flexible, but the relatively large number of parameters often leads to numerical difficulties and unstable drift parameter estimates. In order to introduce sufficient nonlinearity into the model while at the same time obtain closed-form exact transition density, Bu et al. (2011) proposed to use the so-called Reducible Diffusions (RDs). They assumed that their parametric diffusion Y_t is a strictly monotone sufficiently smooth time-independent transformation of an analytically tractable Basic Diffusion (BD) X_t, i.e.

$$Y_t = V(X_t;\theta)$$

where

$$dX_t = \mu_X(X_t;\vartheta)\,dt + \sigma_X(X_t;\vartheta)\,dW_t \tag{4}$$

which is a simpler parametric diffusion depending on $\mu_X(X_t;\vartheta)$ and $\sigma_X^2(X_t;\vartheta)$. In this setup, $V(x;\theta)$ is the transformation function satisfying $\partial V(x,\theta)/\partial x \neq 0$ for all x and θ, where θ is the transformation parameter vector.

The specification of Y_t depends entirely on the choice of X_t and the transformation function V. Hence, Y_t depends on $\psi = (\vartheta, \theta)'$. Since the purpose is to develop models with closed-form exact transition densities, the Vasicek (1977) and the CIR processes are preferred choices for their closed-form transition densities. Bu et al. (2011) noted that for any given BD X_t, the knowledge of the functional form of $\sigma_Y^2(y;\psi)$ will lead to a unique solution of V. Consequently, a feasible strategy is to choose V such that the resulting RD Y_t has a desired diffusion term $\sigma_Y^2(y;\psi)$. Bu et al. (2011) argued a desirable diffusion term is particularly important for financial applications such as volatility modelling and option pricing which rely mostly on the short-run dynamics of the stochastic process.

As discussed above, the CEV specification is parsimonious and flexible. Bu et al. (2011) therefore proposed a nonlinear RD with CEV, where the BD X_t follows the well known CIR process

$$dX_t = (\alpha_0 + \alpha_1 X_t)dt + \sigma X_t^{1/2} dW_t$$

The CIR process is stationary on $(0, +\infty)$ when $\alpha_1 < 0$, $\alpha_0 > 0$ and $2\alpha_0 \geq \sigma^2$. Conditional on $X_{t-\Delta}$, X_t follows the non-central χ^2 distribution with fractional degrees of freedom. They showed that for $\sigma_Y^2(y;\psi) = \sigma^2 y^{2\theta}$, the inverse of the transformation function V can be obtained in closed-form as

$$x = U(y;\theta) = \frac{1}{4}\left(\frac{y^{1-\theta}}{1-\theta}\right)^2$$

The above transformation is strictly monotone on $(0,\infty)$ for $\theta \in (0,\infty)$. It follows from Ito's Lemma that the CIRCEV specification is given by

$$dY_t = \left\{\left[2\alpha_0(1-\theta) + \frac{1}{2}\sigma^2(2\theta-1)\right] Y_t^{2\theta-1} + \frac{\alpha_1 Y_t}{(2-2\theta)}\right\} dt + \sigma Y_t^{\theta} dW_t$$

Bu et al. (2011) showed the CIRCEV specification generates a much stronger pull at high levels of Y_t than the linear drift. It also nests the CIR model ($\theta = 1/2$) and the Ahn and Gao (1999) model ($\theta = 3/2$) as special cases. Clearly, the CIRCEV specification is a general setup which provides not only the nonlinearity in both the drift and diffusion terms but also the extra degrees of freedom in the data-driven choice of θ.

By construction, the transition density can be easily obtained by a change of variable as

$$p_Y(\Delta, y|y_0;\psi) = \frac{y^{1-2\theta}}{2\sqrt{(1-\theta)^2}} c e^{-u-v}\left(\frac{v}{u}\right)^{q/2} I_q\left[2(uv)^{1/2}\right] \tag{5}$$

where

$$c = \frac{2\alpha_1}{\sigma^2\left(e^{\alpha_1\Delta}-1\right)}, u = \frac{ce^{\alpha_1\Delta}}{4}\left(\frac{y_0^{1-\theta}}{1-\theta}\right)^2, v = \frac{c}{4}\left(\frac{y^{1-\theta}}{1-\theta}\right)^2, q = \frac{2\alpha_0}{\sigma^2} - 1$$

and $I_q(\cdot)$ is the modified Bessel function of the first kind of order q. The closed-form transition density allows the users to implement exact ML inference very easily. This is in contrast to the majority of parametric diffusions in the literature (c.f. Durham and Gallant 2002)[4].

3. Regime-Switching Mechanism

3.1. Regime-Switching Diffusion Models

There is a large literature supporting the existence of stochastic regime changes in the evolution of financial and economic variables. Examples include Hamilton (1988), Cai (1994), Gray (1996), Garcia and Perron (1996), and recently Chang et al. (2014). Most recent studies on regime-switching diffusions are by Choi (2009), Goutte and Zou (2013) and Bu et al. (2016).

Following most of the previous works on regime-switching models, we assume that there could be two possible regimes. The dynamics of VIX is therefore described by the following regime-switching continuous-time diffusion process

$$dY_t = \mu_Y \left(Y_t; \psi_{s_t}\right) dt + \sigma_Y \left(Y_t; \psi_{s_t}\right) dW_t$$

where $\mu_Y \left(Y_t; \psi_{s_t}\right)$ and $\sigma_Y \left(Y_t; \psi_{s_t}\right)$ are the regime-dependent drift and diffusion terms and s_t is the regime index. Following the literature, we use low (L) and high (H) regimes to characterize two different economic environments.

Our regime-switching continuous-time framework assumes that the VIX evolves continuously when the economy is in one of the regimes, but allows this continuous-time dynamics to change depending on the state of the economy. Choi (2009) argued that such a regime-switching specification is able to explain the volatility clustering better than a single-regime model. Also, the design of the model allows for two distinct sources of nonlinearity. The first is from the nonlinearity in the regime-dependent diffusion specification, and the second is from the regime-switching dynamics. Consequently, our framework allows us to examine whether the observed nonlinearity in reality should be attributed to the former, or the latter, or both, and hence provide useful guidance for practical applications.

To formally investigate the empirical issues such as nonlinearity, switching components, and endogeneity, we consider two general regime-switching nonlinear diffusion frameworks. The first is the regime-switching NLDCEV (RS-NLDCEV) framework and the second is the regime-switching (RS-CIRCEV) framework. In addition, as a benchmark, the RS-CIR model is also included in our study.

[4] Bu *et al.* (2011) derived the necessary and sufficient conditions for stationarity and unattainability of the boundaries (i.e. 0 and ∞) in finite expected time, as well as the β-mixing property which ensures that the classical asymptotic theory holds for the ML inference based on discretely observed random samples. See Bu *et al.* (2011) for more discussion.

We assume that the regime variable s_t follows a discrete-time first order two-state Markov chain with the following transition matrix

$$P = \begin{pmatrix} p_{LL} & p_{HL} \\ p_{LH} & p_{HH} \end{pmatrix} \tag{6}$$

where we define the transition probabilities as $p_{ij} = P(s_t = j|s_{t-\Delta} = i)$, $i, j = L, H$. Representing stationary distribution of the Markov chain in terms of p_{LL} and p_{HH}, we have

$$\pi_L = \frac{1 - p_{HH}}{2 - p_{LL} + p_{HH}} \text{ and } \pi_H = \frac{1 - p_{LL}}{2 - p_{LL} + p_{HH}}$$

In this study, the transition probabilities p_{LL} and p_{HH} and the parameters of regime-switching diffusion models are estimated jointly by ML based on a Hamilton (1989) type filtering algorithm[5].

3.2. Endogenous Regime-Switching Specification

The above standard regime-switching model assumes constant regime-switching transition probabilities. However, studies including Diebold et al. (1994), Kim et al. (2008), Choi (2009), Chang et al. (2014), and Bu et al. (2016) all reported evidence of endogeneity in regime changes. A convenient way to allow endogenous regime changes in regime-switching models is to specify the transition probabilities as functions of the lagged of the state variable Y_t. In order to capture potential nonlinearity in such a dependence structure, we follow Bu et al. (2016) by considering the following specification for the endogenous transition probabilities

$$p_{LL}(Y_{t-\Delta}) = P(s_t = L|s_{t-\Delta} = L, Y_{t-\Delta}) = \Lambda\left(c_L + \sum_{i=1}^{p} d_{L_i} Y_{t-\Delta}^i\right) \tag{7}$$

$$p_{HH}(Y_{t-\Delta}) = P(s_t = H|s_{t-\Delta} = H, Y_{t-\Delta}) = \Lambda\left(c_H + \sum_{i=1}^{q} d_{H_i} Y_{t-\Delta}^i\right) \tag{8}$$

where Λ is the logistic function. Clearly, if all the coefficients of the powers of $Y_{t-\Delta}$ are jointly zero, the transition probabilities become constant and the endogenous regime-switching model reduces to the exogenous case.

4. Data and Results

4.1. Data

Daily CBOE S&P 500 Volatility Index (VIX) data from January 2, 1990 to March 20, 2015 are considered in this study. *Table 1* provides some summary statistics. There

[5]The Hamilton (1989) algorithm is standard. For space economy, we do not elaborate specific details.

are 6352 observations in our sample. The level of VIX ranges between 9.310 and 80.860 with sample mean 19.921 and standard deviation 7.982. The marginal distribution of VIX shows clear departure from normality with a positive skewness of 2.072 and a large kurtosis of 10.466. The normality is formally rejected by the significance of the enormous Jarque-Bera statistic at 1% significance level.

[Table 1 here]

Time series plots of the VIX series and its first difference are provided in *Figure 1.* We can see that the level of VIX peaked around 1990-1991, 1997-1999, 2002-2003, and 2008-2010. In particular, due to the Global Financial Crisis, the 2008 to 2010 years recorded substantially higher levels of VIX than the rest of the sample period. It is also clear that associated with high levels of VIX are high levels of the volatility of VIX. This feature is consistent with the leverage effect implied by the CEV diffusion specification. Judging from the plots, we can argue that the volatility of VIX behaves quite differently in different time periods. This motivates us to conjecture that regime changes are likely to be an important feature of our data and therefore to analyze diffusion models in regime-switching contexts.

[Figure 1 here]

4.2. Models and Results

For each of our three diffusion specifications, namely the CIR, the NLDCEV, and the CIRCEV, seven distinct models in terms of regime-switching specifications are estimated by ML based on our closed-form transition density functions and the Hamilton-type filtering algorithm. The details of our models and estimation results are summarized in *Table 2, 3, and 4,* respectively. Taking the CIR specification as an example, the seven models are: (i) the single-regime CIR model (CIR), (ii) the general regime-switching CIR model (RS-CIR) with switching drift and diffusion terms, (iii) the drift-switching only RS-CIR model (RS-CIR-1), (iv) the diffusion-switching only RS-CIR model (RS-CIR-2), (v) the endogenous RS-CIR model (ERS-CIR), (vi) the endogenous drift-switching only RS-CIR (ERS-CIR-1), and finally (vii) the endogenous diffusion-switching only RS-CIR (ERS-CIR-2). The same regime-switching specifications for the NLDCEV and the CIRCEV diffusion specifications are created and estimated. We define the low (L) regime as the low VIX volatility regime and the high (H) regime as the high VIX volatility regime. For drift-switching only models, however, we define low (L) and high (H) regimes as the low and high mean reversion regimes, respectively. For all the models, we report ML estimates of model parameters and their standard errors, as well as the maximized log-likelihood (LL), AIC and BIC values.

[Table 2, 3, and 4 here]

4.2.1. Regime-Switching, Endogeneity, and Switching Components

First and foremost, across all three diffusion specifications, the LLs increase substantially when we move from the single-regime model to any of our regime-switching models. However, it is well known that the standard LR test cannot be used to test the presence of two regimes against the null of a single regime, since the parameters related to the second regime are not identified under the null hypothesis. Nevertheless, as suggested by Hamilton and Susmel (1996), Gray (1996), and Choi (2009), the standard LR statistic can still serve informally as a broad indication of the relative performance of competing models in terms of goodness-of-fit to the data. Following this argument, we calculated the LR statistics between the single-regime model and the general (drift and diffusion both switching) exogenous regime-switching model under all three frameworks and reported the results in the first panel of *Table 5*. The LR statistics are 1849.408, 990.93, and 890.664 for the CIR, the NLDCEV, and the CIRCEV frameworks, respectively. These LR statistics would lead to negligible p-values, should the usual Chi-squared distribution with corresponding degrees of freedom be used. Arguably, the improvement in terms of LL is enormous. According to the AIC and BIC values, all our regime-switching models outperformed the single-regime model across all three diffusion specifications. On this basis, it is reasonable for us to carry out our subsequent analysis under the presumption that regime-switching is present in the data.

To test the presence of endogenous regime changes, we choose the maximum lag length of $q = p = 2$ in our estimation[6] of endogenous regime-switching models but only reported the significant lags of Y_t. Our specification is more flexible than most studies where typically only the first lag is included. We therefore expect our model to be more capable of detecting potential, particularly nonlinear, endogeneity. From the second panel of *Table 5*, we can see that across all three diffusion specifications, the assumption of exogenous regime changes is strongly rejected according to the LR statistics calculated between exogenous and endogenous regime-switching models under the same regime-switching specification. It is important to stress that the LR test is a valid test, since the exogenous regime-switching model is strictly nested by the endogenous regime-switching model. Our results confirmed the findings of, for example, Choi (2009), Chang et al. (2014) and Bu et al. (2016), that regime changes in the economy is most likely to be endogenously driven.

[Table 5 here]

The time series plot of VIX and first difference suggest clear evidence of changing volatilities of VIX over different periods of time. We therefore conjecture that models with a regime-switching diffusion component may capture the VIX dynamics

[6] Adding more lags did not improve the LL significantly.

better than models without a switching diffusion component. Meanwhile, it is well known that the drift and diffusion terms play very different roles in the dynamics of continuous-time diffusions. Specifically, the drift determines the long-run behavior (e.g. mean reversion), whereas the diffusion term drives the local (i.e. short-run) activities. Consequently, it usually requires a long span of data to estimate the drift satisfactorily (c.f. Bandi and Phillips 2003). For our regime-switching models, there are two drift terms to be identified, while at the same time only certain segments of the sample contain information about the long-run behaviors of the two regime-dependent diffusions. This is likely to reduce the model's ability to identify the drift term accurately and thus weaken its power of capturing certain features of the VIX dynamics. In contrast, the estimation of regime-dependent diffusion terms is less affected as the diffusion term captures the local feature of the data. For these reasons, we expect the diffusion-switching only models to fit the data better than the drift-switching only models.

We resort again to the LL, information criteria and LR statistics to test models with one switching component against the models with both components switching. First, as can be already seen from *Table 2, 3, and 4*, our conjecture that diffusion-switching only models perform better than drift-switching only models is confirmed by the LL, AIC and BIC values. Across all three frameworks, diffusion-switching only models are uniformly preferred than drift-switching only models under the same regime-switching specifications. Second, according to the LR test results reported in the third and fourth panels of *Table 5*, drift-switching only models and diffusion-switching only models are both rejected against the general both-switching models significance the tests. This suggests that regime-switching effects in the evolution of VIX are likely to affect both the long-run behavior (i.e. central tendency) and the short-run behavior (i.e. stochastic volatility) of the VIX dynamics. Meanwhile, this also confirms that our general regime-switching models, despite the relatively large number of parameters, did not overfit the data.

4.2.2. Nonlinearity

The nonlinearity in the drift and diffusion terms is one of the most important issues in diffusion modelling. First, there is an obvious trade-off between the flexibility and the tractability of the diffusion models. For instance, linear (e.g. affine) models are best known for their analytical tractability because of their closed-form conditional characteristic functions. Consequently, ML inference and financial applications such as derivative pricing become particularly convenient. In contrast, nonlinear models usually do not have tractable transition density or conditional characteristic functions[7]. Therefore, the implementation of these models are generally more complicated. On the other hand, while the nonlinearity in the diffusion term is found conclusively sig-

[7] To our best knowledge, the only class of nonlinear diffusions with closed-form transition densities and conditional characteristic functions are the reducible (transformation-based) diffusions. See Bu *et al.* (2011) for more discussion.

nificant in many financial applications, evidence of nonlinearity in the drift is largely inconclusive (c.f. Choi 2009). This leads to the debate whether a nonlinear drift is required in certain applications at the cost of losing tractability. To our best knowledge, few studies addressed this issue in VIX modelling, particularly in the context of regime-switching.

We first consider testing nonlinearity under the NLDCEV framework. The nonlinearity in the drift and the diffusion terms of the NLDCEV specification is determined by α_{-1}, α_2 and θ. For example, the process reduces to the CIR process when $\alpha_{-1} = \alpha_2 = 0$ and $\theta = 1/2$. It follows naturally that it is possible to test nonlinearity by nesting the CIR specification into the more general NLDCEV specification and examine the LR statistic. We then proceed to calculate the relevant LR between the CIR and NLDCEV models across all regime-switching specifications and report them in the upper panel of *Table 6*. Across all cases, the LR tests are significant at least at 1% level, strongly rejecting all models based on the linear CIR specification in favor of the more general NLDCEV specification for any given regime-switching specification.

[Table 6 here]

Meanwhile, testing nonlinearity under the CIRCEV framework is considerably easier. Since the CIRCEV process is a smooth transformation of the CIR process, the nonlinearity in the drift and the diffusion terms is determined by the value of the transformation parameter θ alone. Specifically, the CIRCEV specification reduces to the CIR specification when $\theta = 1/2$. Results of the relevant LR between the CIR and CIRCEV models across all regime-switching specifications are reported in the lower panel of *Table 6*. Similar to the NLDCEV case, these tests strongly reject all models based on the CIR process in favor of models based on the CIRCEV process across all regime-switching specifications.

The strong rejection of the linear CIR specification against the more general NLDCEV and CIRCEV specifications should not be taken as a surprise, since the dynamics of the former is known to be excessively simple and empirically rejected in many different contexts (e.g. Choi 2009, Bu et al. 2011) against nonlinear alternatives. However, the most important implication of our results is that the obvious nonlinearity in the dynamics of VIX cannot be explained by regime-switching effects alone.

4.2.3. Comparison Between Nonlinear Specifications

Finally, we examine the relative performance between models based on the NLDCEV specification and models based on the CIRCEV specification. As discussed, both nonlinear specifications are capable of generating desirable shapes of the drift and diffusion functions consistent with many empirical data. While they share the same diffusion function, they differ in their drift specifications. The obvious advantage of the CIRCEV specification is that the reducibility ensures that its exact transition density can be evaluated in closed-form, while for the NLDCEV specification the transition density has to be approximated. The NLDCEV also has two more parameters than the CIRCEV. Clearly, the CIRCEV is substantially more tractable and parsimonious. For this reason, it would be very useful, from a practical point of view, to compare the empirical performance of the two specifications in the context of our regime-switching specifications to see if the effort of density approximation and additional parameters are worthwhile.

We begin with an informal comparison based on the maximized LLs and the two information criteria. First, the maximized LLs of CIRCEV-based models are substantially larger than those of NLDCEV-based models uniformly across all seven regime-switching specifications. We also report the standard LR statistics and their statistical significance in the upper panel of *Table 7*. Clearly, the LRs are so large, significant at 1% level, that the p-values would have been practically zero, if the usual Chi-squared distribution with corresponding degrees of freedom should be used. Although such a test would not have been strictly valid due to the overlapping nature of the two models, the huge difference in the LLs is indicative of the fact that CIRCEV-based models provide much better fit to the data than NLDCEV-based models across all regime-switching specifications. Meanwhile, a comparison based on the information criteria is a more legitimate exercise. Since the NLDCEV has two more parameters than the CIRCEV, the information criteria penalize the NLDCEV-based models even further. Therefore, at least in loose terms the CIRCEV specification outperforms the NLDCEV specification uniformly across all our regime-switching specifications.

[Table 7 here]

To investigate this issue formally, the main difficulty lies in the overlapping nature of the two models[8]. In addition, the two models are dynamic. Therefore, neither the Vuong (1989) test, nor the Rivers and Vuong (2002) test is suitable. We therefore resort to the recent test developed by Marcellino and Rossi (2008), henceforth denoted as the MR test, for comparing overlapping competing nonlinear dynamic models. The null hypothesis of the MR test is that the models are equally close to the true DGP according to the Kullback and Leibler (1951) information criterion. The alternative is that one model is closer to the true DGP. More importantly, their test is valid whether or not the models are correctly specified and suitable for a variety of estimation methods. Since we estimated our models by ML, the MR test statistic is a normalized version of the standard LR statistic. Under the null hypothesis, the test statistic has either a normal or a mixture of Chi-squared asymptotic distribution depending on the speed of convergence of the variance of the test statistic. See Marcellino and Rossi (2008) for more details.

[8] This can be seen from the fact that the CIR specification is nested in both the NLDCEV and the CIRCEV specifications.

The test statistics and their significance are reported in the lower panel of *Table 7*. Not surprisingly, all seven null hypotheses are strongly rejected. This is valid and convincing evidence to conclude that the CIRCEV specification does outperform the NLDCEV specification at least in the context of our data and regime-switching specifications. This particular conclusion may carry a significant implication for the empirical practice of modelling nonlinear continuous-time diffusions, particularly when regime switching is allowed. Clearly, not only does the CIRCEV has closed-form exact transition density, which complete avoids the discretization or approximation bias (c.f. Hsiao and Semmler 2011), but it also provide important functional flexibility even better than the more heavily parameterized NLDCEV.

4.3. Conclusion

We studied model specification in regime-switching continuous-time diffusions for market volatility, using the VIX which the widely accepted fear gauge as our proxy. The frameworks under which our analysis was carried out are general and flexible enough for us to examine a number of important empirical issues such as regime-switching, endogeneity, switching components, nonlinearity, as well as to compare the relative performance of the two closely related yet quite distinct nonlinear diffusion specifications, namely the NLDCEV and the CIRCEV specifications, on the basis of formal and informal criteria.

At least in the context of our diffusion and regime-switching specifications, we found strong evidence of regime changes in the dynamics of market volatility. In particular, our results suggest that such a regime-switching mechanism is primarily characterized by the changing volatilities of the market volatility and most likely endogenously driven by the level of market volatility itself. Moreover, regime-dependent nonlinearity remained to be a crucial feature of the underlying dynamics, as the regime-switching effects alone cannot account for the complicated nonlinear dynamics sufficiently. Consequently, we conclude that the most suitable models for modelling market volatility such as the VIX should have an endogenous regime-switching mechanism and at least suitably specified regime-dependent nonlinear diffusion functions. Compared to the diffusion term, the role of the drift term is likely to be of secondary importance. For these reasons, nonlinear diffusion specifications with desirable tractability, such as the reducible diffusions, can potentially play a very important role in this type of situations.

Further research can include conducting similar studies for modelling other financial variables such as the short-term interest rates and exchange rates (c.f. Choi 2009, Goutte and Zou 2013), or extending the current frameworks to models with jumps (c.f. Chevallier and Goutte 2014), or considering pricing financial derivatives based on regime-switching affine reducible continuous-time models (c.f. Amengual and Xiu 2012, Mencia and Sentana 2013).

References

Ahn, D.H., and Gao, B. (1999): A Parametric Non-linear Model of Term Structure Dynamics. *Review of Financial Studies*, 12, 721-762.

Aït-Sahalia, Y. (1996a): Nonparametric Pricing of Interest Rate Derivatives. *Econometrica*, 64, 527-560.

Aït-Sahalia, Y. (1996b): Testing Continuous-Time Models of the Spot Interest Rate. *Review of Financial Studies*, 9, 385-426.

Aït-Sahalia, Y. (1999): Transition Densities for Interest Rate and Other Nonlinear Diffusions. *Journal of Finance*, 54, 1361-1395.

Aït-Sahalia, Y. (2002): Maximum Likelihood Estimation of Discretely Sampled Diffusions: A Closed-Form Approximation Approach. *Econometrica*, 70(1), 223-262.

Amengual, D., and Xiu, D. (2012): Delving into Risk Premia: Reconciling Evidence from the S&P 500 and VIX Derivatives. *Technical Report*, University of Chicago Booth School of Business.

Bandi. F.M., and Phillips, P.C.B. (2003): Fully Nonparametric Estimation of Scalar Diffusion Models. *Econometrica*, 71(1), 241-283.

Bai, J., and Perron, P. (1998): Estimating and Testing Linear Models with Multiple Structural Changes. *Econometrica*, 66. 47-68.

Black, F., and Scholes, M. (1973): The Pricing of Options and Corporate Liabilities, *Journal of Political Economy*, 81, 637-654.

Bu, R., Cheng, J., and Hadri, K. (2016): Reducible Diffusions with Time-Varying Transformations with Application to Short-Term Interest Rates. *Economic Modelling*, 52, 266-277.

Bu, R., Hadri, K., and Kristensen, D. (2015): Transformed Diffusions and Copulas: Identification and Estimation. *Working Paper.*

Bu, R., Giet, L., Hadri, K., and Lubrano, M. (2011): Modeling Multivariate Interest Rates Using Time-Varying Copulas and Reducible Nonlinear Stochastic Differential Equations. *Journal of Financial Econometrics*, 9(1), 198-236.

Cai, J. (1994): A Markov Model of Switching-Regime ARCH. *Journal of Business and Economic Statistics*, 12(3), 309-316.

Chan, K.C., Karolyi, A., Longstaff, F., and Sanders, A. (1992): An Empirical Comparison of Alternative Models of the Short-Term Interest Rate. *Journal of Finance*, 47, 1209-1227.

Chang, Y., Choi, Y., and Park, J. (2014): Regime Switching Model with Endogenous Autoregressive Latent Factor. *Working Paper.*

Chen, J.H., and Huang, Y.F. (2014): Long Memory and Structural Breaks in Modelling the Volatility Dynamics of VIX-ETFS. *International Journal of Business, Economics and Law*, 4(1), 54-63.

Chevallier, J., and Goutte S. (2015): Detecting Jumps and Regime Switches in International Stock Markets Returns. *Applied Economics Letters*, 22(13), 1011-1019.

Choi, S. (2009): Regime-Switching Univariate Diffusion Models of the Short-Term Interest Rate. *Studies in Nonlinear Dynamics and Econometrics*, 13(1), Article 4.

Conley, T.L., Hansen, L., Luttmer, E., and Scheinkman, J. (1997): Short Term Interest Rates as Subordinated Diffusion. *Review of Financial Studies*, 10, 525-578.

Cox, J., Ingersoll, J., and Ross, S. (1985): In Intertemporal General Equilibrium Model of Asset Prices. *Econometrica*, 53, 363-384.

Diebold, F., Lee, J., and Weinbach, G. (1994): Regime Switching with Time-Varying Transition Probabilities. In C. Hargreaves (Ed.), *Nonstationary Time Series Analysis and Cointegration*, 283-302. Oxford University Press.

Duffee, D., and Kan, R. (1996): A Yield Factor Model of Interest Rates. *Mathematical Finance*, 6, 379-406.

Durham, G., and Gallant, R. (2002): Numerical Techniques for Maximum Likelihood Estimation of Continuous-Time Diffusion Processes. *Journal of Business and Economic Statistics*, 20(3) 297-338.

Durham, G. (2003): Likelihood-Based Specification Analysis of Continuous-Time Models of the Short-Term Interest Rate. *Journal of Financial Economics*, 70, 463-487.

Gallant, A.R., and Tauchen, G. (1998): Reprojecting Partially Observed Systems with Application to Interest Rate Diffusions. *Journal of the American Statistical Association*, 93(441), 10-24.

Garcia, R., and Perron, P. (1996): An Analysis of the Real Interest Rate under Regime Shifts. *Review of Economics and Statistics*, 78(1), 111-125.

Gray, S.F. (1996): Modeling the Conditional Distribution of Interest Rates as a Regime-Switching Process. *Journal of Financial Economics*, 42, 27-62.

Goutte, S., and Zou, B. (2013): Continuous Time Regime Switching Model Applied to Foreign Exchange Rate. *Mathematical. Finance Letters*, 2013:8.

Guo, W., and Wohar, M.E. (2006): Identifying Regime Changes in Market Volatility. *Journal of Financial Research*, 29(1), 79-93.

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 Nonstationary Time Series. *Econometrica*, 57, 357-384.

Hamilton, J., and Susmel, R. (1994): Autoregressive Conditional Heteroskedasticity and Changes in Regime. *Journal of Econometrics*, 64, 307-333.

Hsiao, C., and Semmler, W. (2011): Continuous and Discrete Time Modeling of Short-Term Interest Rates, in: Gregoriou, G., and Pascalau, R. (Ed.), Financial Econometrics Modeling: Derivatives Pricing, Hedge Funds and Term Structure Models, Palgrave, 163-187.

Jiang, G., and Knight, J. (1997): A Nonparametric Approach to the Estimation of Diffusion Processes with an Application to a Short-Term Interest Rate Model. *Econometric Theory*, 13, 615-645.

Kim, C.J., Piger, J., and Startz, R. (2008): Estimation of Markov Regime-Switching Regression Models with Endogenous Switching. *Journal of Econometrics*, 143, 263-273.

Kullback, S., and Leibler, R.A. (1951): On Information and Sufficiency. *Annals of Mathematical Statistics*, 22(1), 79–86.

Kristensen, D. (2010). Pseudo-Maximum Likelihood Estimation in Two Classes of Semiparametric Diffusion Models. *Journal of Econometrics*, 156, 239-259.

Marcellino, M., and Rossi, B. (2008): Model Selection for Nested and Overlapping Nonlinear, Dynamic and Possibly Mis-specified Models. *Oxford Bulletin of Economics and Statistics*, 70(1), 867-893.

Mencia, J., and Sentana, E. (2013): Valuation of VIX derivatives. *Journal of Financial Economics*, 108(2), 367–391.

Merton, R.C. (1973). Theory of Rational Option Pricing. *Bell Journal of Economics and Management Science*, 4, 141-183.

Papanicolaou, A., and Sircar, R. (2014): A Regime-Switching Heston Model for VIX and S&P 500 Implied Volatilities. *Quantitative Finance*, 14(10), 1811-1827.

Rivers, D., and Vuong, Q. (2002): Model selection tests for nonlinear dynamic models. *Econometrics Journal*, 5, 1–39.

Stanton, R. (1997). A Nonparametric Model of Term Structure Dynamics and the Market Price of Interest Rate Risk. *Journal of Finance*, 52, 1973-2002.

Song, Z., and Xiu, D. (2014): A Tale of Two Option Markets: State-Price Densities Implied from S&P 500 and VIX Option Prices. *Journal of Econometrics*, Forthcoming.

Vasicek, O. (1977): An Equilibrium Characterization of the Term Structure. *Journal of Financial Economics*, 5, 177-188.

Vuong, Q. (1989): Likelihood Ratio Tests for Model Selection and Nonnested Hypotheses. *Econometrica*, 57, 307–333.

Whaley, R.E. (2000): The Investor Fear Gauge. *Journal of Portfolio Management*, 26(3), 12-17.

Table 1: Descriptive statistics of VIX data

Sample Period	Jan 2, 1990 - Mar 20, 2015
Sample Size	6352
Mean	19.921
Min	9.310
Max	80.860
Std Dev.	7.982
Skewness	2.072
Kurtosis	10.466
Jarque-Bera Statistic	19298.001

Table 2: ML estimation results for RS-CIR framework

	CIR	RS-CIR	RS-CIR-1	RS-CIR-2	ERS-CIR	ERS-CIR-1	ERS-CIR-2
$\alpha_{0,L}$	95.047	85.493	43.307	75.262	81.492	32.652	67.390
	(11.634)	(10.712)	(4.555)	(9.455)	(11.734)	(5.754)	(9.675)
$\alpha_{0,H}$		313.941	500.092		349.779	579.605	
		(41.830)	(28.655)		(43.022)	(33.082)	
$\alpha_{1,L}$	-4.779	-5.941	-0.000	-4.802	-5.785	-0.041	-4.313
	(0.619)	(0.659)	(0.000)	(0.557)	(0.749)	(0.039)	(0.563)
$\alpha_{1,H}$		−10.399	-35.196		-11.937	-44.009	
		(1.719)	(1.536)		(1.682)	(2.550)	
σ_L^2	21.635	7.854	18.637	8.082	7.078	19.455	7.160
	(0.988)	(0.186)	(0.371)	(0.189)	(0.272)	(0.377)	(0.281)
σ_H^2		55.893		57.134	54.434		54.075
		(1.755)		(1.784)	(2.505)		(2.484)
P_{LL}		0.959	0.767	0.962			
		(0.012)	(0.014)	(0.012)			
P_{HH}		0.898	0.138	0.901			
		(0.012)	(0.047)	(0.013)			
c_L					1.307	13.831	1.271
					(3.170)	(3.087)	(2.931)
c_H					-18.935	11.224	-21.529
					(7.513)	(5.886)	(7.883)
d_{L_1}						-0.735	
						(0.225)	
d_{L_2}						0.014	
						(0.005)	
d_{H_1}					2.614		3.008
					(1.039)		(1.110)
d_{H_2}					-0.110		-0.128
					(0.046)		(0.050)
LL	-10474.402	-9549.698	-10317.912	-9589.703	-9490.337	-10260.062	-9530.299
AIC	20954.804	19115.396	20649.824	19191.406	19000.674	20538.124	19076.598
BIC	20975.074	19169.448	20697.120	19231.945	19068.239	20598.933	19130.650

Table 3: ML estimation results for RS-NLDCEV framework

	NLDCEV	RS-NLDCEV	RS-NLDCEV-1	RS-NLDCEV-2	ERS-NLDCEV	ERS-NLDCEV-1	ERS-NLDCEV-2
$\alpha_{-1,L}$	826.013	4364.769	2379.952	2517.836	4364.423	2379.005	2516.015
	(960.884)	(1508.870)	(865.087)	(955.888)	(1427.770)	(1771.127)	(1437.977)
$\alpha_{-1,H}$		140.329	1449.683		143.820	1448.998	
		(302.001)	(1736.496)		(728.301)	(2707.071)	
$\alpha_{0,L}$	-86.565	-675.427	-661.474	-364.410	-680.209	-660.996	-367.266
	(156.901)	(252.835)	(152.410)	(153.496)	(243.879)	(400.619)	(235.001)
$\alpha_{0,H}$		177.392	81.185		201.721	80.997	
		(74.038)	(153.675)		(132.371)	(238.894)	
$\alpha_{1,L}$	5.261	33.767	70.242	17.968	33.871	70.036	18.175
	(7.510)	(13.212)	(14.176)	(7.698)	(13.139)	(26.718)	(11.846)
$\alpha_{1,H}$		-4.924	-5.734		-6.431	-5.825	
		(3.748)	(3.071)		(7.201)	(5.011)	
$\alpha_{2,L}$	–0.142	-0.581	-2.441	-0.323	-0.568	-2.434	-0.323
	(0.099)	(0.214)	(0.482)	(0.121)	(0.223)	(0.492)	(0.183)
$\alpha_{2,H}$		-0.000	-0.000		-0.018	-0.004	
		(0.000)	(0.000)		(0.112)	(0.001)	
σ_L^2	0.151	0.079	0.156	0.070	0.044	0.156	0.039
	(0.022)	(0.019)	(0.023)	(0.016)	(0.010)	(0.023)	(0.009)
σ_H^2		1.849		1.599	1.799		1.518
		(0.887)		(0.062)	(0.791)		(0.628)
θ_L	1.314	1.325	1.310	1.349	1.422	1.311	1.444
	(0.024)	(0.041)	(0.025)	(0.020)	(0.039)	(0.025)	(0.038)
θ_H		1.075		1.103	1.079		1.105
		(0.078)		(0.039)	(0.072)		(0.066)
P_{LL}		0.965	0.999	0.969			
		(0.013)	(0.011)	(0.013)			
P_{HH}		0.843	0.999	0.849			
		(0.021)	(0.011)	(0.020)			
c_L					6.027	6.335	1.444
					(3.005)	(1.664)	(3.024)
c_H					-9.704	5.840	1.105
					(2.407)	(5.277)	(3.055)
d_{L_1}						1.141	3.081
						(0.482)	(0.411)
d_{L_2}					0.043		-0.251
					(0.020)		(0.016)
d_{H_1}					1.086		-10.164
					(0.260)		(0.360)
d_{H_2}					-0.029		1.188
					(0.009)		(0.013)
LL	-9837.452	-9341.993	-9810.662	-9369.349	-9315.848	-9807.289	-9341.706
AIC	19686.904	18711.986	19645.324	18758.698	18665.696	19640.578	18711.412
BIC	19727.443	18806.577	19726.402	18826.263	18780.557	19728.413	18806.003

Table 4: ML estimation results for RS-CIRCEV framework

	CIRCEV	RS-CIRCEV	RS-CIRCEV-1	RS-CIRCEV-2	ERS-CIRCEV	ERS-CIRCEV-1	ERS-CIRCEV-2
$\alpha_{0,L}$	0.628	0.654	0.745	0.684	0.198	0.916	0.189
	(0.196)	(0.327)	(0.264)	(0.305)	(0.102)	(0.380)	(0.073)
$\alpha_{0,H}$		6.645	0.193		6.519	1.040	
		(6.357)	(0.333)		(8.034)	(1.401)	
$\alpha_{1,L}$	-3.844	-2.184	-3.242	-2.785	-1.631	-3.515	-2.449
	(0.555)	(0.529)	(0.528)	(0.481)	(0.554)	(0.592)	(0.478)
$\alpha_{1,H}$		–11.220	-39.671		-12.028	-42.749	
		(2.351)	(4.246)		(2.372)	(6.268)	
σ_L^2	0.091	0.062	0.081	0.060	0.035	0.088	0.030
	(0.014)	(0.014)	(0.014)	(0.013)	(0.009)	(0.017)	(0.006)
σ_H^2		0.538		0.298	0.502		0.146
		(0.228)		(0.066)	(0.278)		(0.032)
θ_L	1.396	1.361	1.380	1.369	1.457	1.367	1.484
	(0.026)	(0.039)	(0.029)	(0.037)	(0.042)	(0.032)	(0.034)
θ_H		1.270			1.280		
		(0.071)			(0.090)		
P_{LL}		0.963	0.984	0.966			
		(0.013)	(0.036)	(0.021)			
P_{HH}		0.847	0.037	0.859			
		(0.021)	(0.021)	(0.017)			
c_L					3.894	1.918	2.260
					(2.760)	(1.133)	(1.677)
c_H					-10.067	-3.871	-1.891
					(2.174)	(1.889)	(0.364)
d_{L_1}						0.255	
						(0.126)	
d_{L_2}					0.034	-0.008	0.009
					(0.016)	(0.004)	(0.005)
d_{H_1}					1.146		0.270
					(0.228)		(0.051)
d_{H_2}					-0.031		-0.004
					(0.007)		(0.002)
LL	-9777.305	-9331.973	-9546.243	-9348.018	-9304.674	-9540.546	-9326.202
AIC	19562.610	18683.946	19108.486	18710.036	18635.348	19101.092	18672.404
BIC	19589.636	18751.511	19162.538	18757.332	18723.183	19168.657	18739.969

Table 5: LR test results for regime-switching, endogeneity and switching components

H_0(one-regime)	vs	H_1(two-regimes)	LR Statisitc	p-value
CIR	vs	RS-CIR	1849.40	0.000
NLDCEV	vs	RS-NLDCEV	1303.68	0.000
CIRCEV	vs	RS-CIRCEV	990.93	0.000

H_0(exogenous)	vs	H_1(endogeneous)	LR Statisitc	p-value
RS-CIR	vs	ERS-CIR	118.722	0.000
RS-CIR-1	vs	ERS-CIR-1	115.700	0.000
RS-CIR-2	vs	ERS-CIR-2	118.808	0.000
RS-NLDCEV	vs	ERS-NLDCEV	52.280	0.000
RS-NLDCEV-1	vs	ERS-NLDCEV-1	6.746	0.009
RS-NLDCEV-2	vs	ERS-NLDCEV-2	55.286	0.000
RS-CIRCEV	vs	ERS-CIRCEV	54.598	0.000
RS-CIRCEV-1	vs	ERS-CIRCEV-1	11.394	0.004
RS-CIRCEV-2	vs	ERS-CIRCEV-2	43.632	0.000

H_0(drift)	vs	H_1(both)	LR Statisitc	p-value
RS-CIR-1	vs	RS-CIR	1536.428	0.000
RS-NLDCEV-1	vs	RS-NLDCEV	937.338	0.000
RS-CIRCEV-1	vs	RS-CIRCEV	428.540	0.000
ERS-CIR-1	vs	ERS-CIR	1539.450	0.000
ERS-NLDCEV-1	vs	ERS-NLDCEV	982.882	0.000
ERS-CIRCEV-1	vs	ERS-CIRCEV	471.744	0.000

H_0(diffusion)	vs	H_1(both)	LR Statisitc	p-value
RS-CIR-2	vs	RS-CIR	80.010	0.000
RS-NLDCEV-2	vs	RS-NLDCEV	54.712	0.000
RS-CIRCEV-2	vs	RS-CIRCEV	32.090	0.000
ERS-CIR-2	vs	ERS-CIR	79.924	0.000
ERS-NLDCEV-2	vs	ERS-NLDCEV	51.716	0.000
ERS-CIRCEV-2	vs	ERS-CIRCEV	43.056	0.000

Table 6: LR test results for nonlinearity

H_0(linear)	vs	H_1(nonlinear)	LR statistic	p-value
CIR	vs	NLDCEV	1273.900	0.000
RS-CIR	vs	RS-NLDCEV	415.410	0.000
RS-CIR-1	vs	RS-NLDCEV-1	1014.500	0.000
RS-CIR-2	vs	RS-NLDCEV-2	440.708	0.000
ERS-CIR	vs	ERS-NLDCEV	348.978	0.000
ERS-CIR-1	vs	ERS-NLDCEV-1	905.546	0.000
ERS-CIR-2	vs	ERS-NLDCEV-2	377.186	0.000

H_0(linear)	vs	H_1(nonlinear)	LR statistic	p-value
CIR	vs	CIRCEV	1394.194	0.000
RS-CIR	vs	RS-CIRCEV	435.450	0.000
RS-CIR-1	vs	RS-CIRCEV-1	1543.338	0.000
RS-CIR-2	vs	RS-CIRCEV-2	483.370	0.000
ERS-CIR	vs	ERS-CIRCEV	371.326	0.000
ERS-CIR-1	vs	ERS-CIRCEV-1	1439.032	0.000
ERS-CIR-2	vs	ERS-CIRCEV-2	408.194	0.000

Table 7: **LR and MR test results for comparing models**

H_0(equal)	vs	H_1(unequal)	LR statistic	p-value
NLDCEV	vs	CIRCEV	120.294	0.000
RS-NLDCEV	vs	RS-CIRCEV	20.040	0.001
RS-NLDCEV-1	vs	RS-CIRCEV-1	528.838	0.000
RS-NLDCEV-2	vs	RS-CIRCEV-2	42.662	0.000
ERS-NLDCEV	vs	ERS-CIRCEV	22.348	0.000
ERS-NLDCEV-1	vs	ERS-CIRCEV-1	533.486	0.000
ERS-NLDCEV-2	vs	ERS-CIRCEV-2	31.008	0.000

H_0(equal)	vs	H_1(unequal)	MR statistic	p-value
NLDCEV	vs	CIRCEV	2.959	0.002
RS-NLDCEV	vs	RS-CIRCEV	1.846	0.032
RS-NLDCEV-1	vs	RS-CIRCEV-1	4.496	0.000
RS-NLDCEV-2	vs	RS-CIRCEV-2	3.708	0.000
ERS-NLDCEV	vs	ERS-CIRCEV	2.059	0.019
ERS-NLDCEV-1	vs	ERS-CIRCEV-1	4.419	0.000
ERS-NLDCEV-2	vs	ERS-CIRCEV-2	2.006	0.023

Figure 1: Time series and differenced series of VIX data