# Reducible Diffusions with Time-Varying Transformations with Application to Short-Term Interest Rates

Ruijun Bu[*]
University of Liverpool

Jie Cheng
Xi'an Jiaotong-Liverpool University

Kaddour Hadri
Queen's University Belfast

[*]Corresponding author: Management School, Chatham Street, Liverpool, L69 7ZH, UK, Tel: +44-151-795-3122, Fax: +44-151-795-3004, Email: RuijunBu@liv.ac.uk (Ruijun Bu).

# Reducible Diffusions with Time-Varying Transformations with Application to Short-Term Interest Rates

## Abstract

Reducible diffusions (RDs) are nonlinear transformations of analytically solvable Basic Diffusions (BDs). Therefore, they are constructed to be analytically tractable and flexible diffusion processes. Existing literature on RDs has mostly focused on time-homogeneous transformations, which to a significant extent fail to explore the full potential of RDs from both theoretical and practical point of views. In this paper, we propose flexible and economically justifiable time-variations to the transformations of RDs. Concentrating on the Constant Elasticity Variance (CEV) RDs, we consider nonlinear dynamics for our time-varying transformations with both deterministic and stochastic designs. Such time-variations can greatly enhance the flexibility of RDs while maintain sufficient tractability of the resulting models. Our approach also enjoys the benefits of classical inferential techniques as much as the advocated time-varying nonlinear dynamics. Our application to UK and US short-term interest rates suggests that from an empirical point of view time-varying transformations are highly relevant and statistically significant.

# 1. Introduction

Since the seminal work of Merton (1973), continuous-time diffusion models have proved to be extremely useful in modelling financial and economic dynamics. They have been frequently applied to research in consumption, savings and investment problems, contingent claim pricing, asset return dynamics and so on. In particular, probably more models have been put forward to explain the behavior of short-term interest rates (short-rates) than for any other issue in finance (cf. Chan et al. 1992).

The basic dynamics for a univariate continuous-time diffusion $\{Y_t, t \geq 0\}$ is described by the following Stochastic Differential Equation (SDE), also known as the Ito's diffusion:

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

where $\mu_Y (y)$ and $\sigma_Y^2 (y)$ are the instantaneous drift and diffusion functions respectively, and $\{W_t, t \geq 0\}$ is a standard Brownian motion. Parametric diffusions, which form the majority in the literature, assume that $\mu_Y$ and $\sigma_Y^2$ are known functions up to an unknown parameter vector $\phi$, i.e. $\mu_Y (y) = \mu_Y (y, \phi)$ and $\sigma_Y^2 (y) = \sigma_Y^2 (y, \phi)$. Well known examples of parametric diffusions in finance include Merton (1973), Brennan and Schwartz (1979), Vasicek (1977), Cox (1975), Dothan (1978), Cox et al. (1980, 1985), Courtadon (1982), Constantinides and Ingersoll (1984), Constantinides (1992), Duffie and Kan (1996), Aït-Sahalia (1996b), Conley et al. (1997), Ahn and Gao (1999) (AG), Bu et al. (2011) and so on. Nonparametric and semiparametric approaches which deviate from the full parametric assumptions have also been proposed in the literature for their functional flexibility. Notable examples include Aït-Sahalia (1996a), Stanton (1997), Jiang and Knight (1997), Kristensen (2010), and most recently Bu et al. (2014), etc.

From an econometric point of view, parametric diffusions often provide a more intuitive and convenient way to specify the dynamics of the state variable. In the meantime, it is also convenient to apply classical inferential techniques such as Maximum Likelihood (ML) and Method of Moments as long as the likelihood function or certain moment functions can be evaluated effectively. In this regard, inference for nonparametric or semiparametric diffusions can be significantly more complicated and inefficient (cf. Kristensen 2010). Thus, from a practical point of view, parametric diffusions are much more abundant and widely used than nonparametric or semiparametric diffusions in empirical applications.

Consequently, a great deal of effort has been spent searching for efficient ways to estimate parametric diffusions. ML is typically the method of choice for its proclaimed efficiency gain. Nevertheless, diffusion models are specified in continuous time, but empirical data are always sampled at discrete-time intervals. Little can be said about the implications of the dynamics in equation (1) for longer time intervals. In finance, Black and Scholes (1973), Vasicek (1977), Cox et al. (1985), Ahn and Gao (1999) are the few rare cases where the discrete-time transition Probability Density Function (PDF) is known in closed form. However, substantial nonlinearity beyond the assumptions of these cases has been

3

documented in the literature. In the context of short-rate modelling, Aït-Sahalia (1996b) for example concluded that the majority of existing parametric diffusion models were rejected by his data. This then became the motivation behind his well known Aït-Sahalia (1996b) general parametric specification[1]. Meanwhile, in a fully nonparametric setting, Stanton (1997) also observed strong nonlinearity in diffusion models for the short-rate series. The difficulty of almost all nonlinear diffusions are two folded. On one hand, they normally have no closed-form transition PDFs. Hence, considerable energy has been employed in developing various density approximation schemes. However, a price has to be paid for approximation errors and computational burden (cf. Durham and Gallant 2002). On the other hand, some parameters of highly nonlinear models can sometimes be hard to identify from the data (cf. Elerian et al. 2001). Therefore, the problem of flexible modelling and efficient estimation of nonlinear continuous-time diffusions remains to be an important issue in practice.

In view of these difficulties, Bu et al. (2011) proposed a novel approach for modeling diffusions using Reducible SDEs. Reducible Diffusions (RDs) are defined by Kloeden and Platen (1992) as monotone transformations of analytically tractable Basic Diffusions (BDs) that have closed-form solutions. Since RDs are usually constructed by nonlinear transformations, they are potentially more flexible to capture complex dynamics of stochastic processes but at the same time possess desirable analytical tractability inherited from tractable BDs. Bu et al. (2011) considered two classes of RDs. The first class are diffusions transformed from the Ornstein-Uhlenbeck (OU) process (cf. Vasicek 1977) and the second are transformations of the square-root (CIR) process (cf. Cox et al. 1985). Since the OU and the CIR processes have renowned analytical tractability, both OU-reducible diffusions (OU-RDs) and CIR-reducible diffusions (CIR-RDs) have similar tractability. In the context of short-rate modelling, they investigated RDs with Constant Elasticity Variance (CEV), i.e. $\sigma_Y^2(y) = \sigma_0^2 y^{2\theta}$, which they named as OU-CEV and CIR-CEV RDs respectively. They showed that OU-CEV and CIR-CEV RDs are power $\theta$ functions of the OU and the CIR processes respectively and nest many known parametric models that have exact closed-form transition PDFs.

Modelling with nonlinear RDs has a number of advantages. Firstly, since RDs are transformations of BDs, additional (often substantial) flexibility can be achieved by specifying suitable nonlinear transformations. Secondly, the most important property of RDs is that their transition PDFs can be expressed in closed form via a transformation of the closed-form transition PDFs of BDs. Thus, the likelihood function for discretely observed samples can be evaluated and then optimized efficiently and standard likelihood-based inferential techniques can be used conveniently. Thirdly, the conditional Cumulative Distribution Functions (CDFs) of RDs, also known as the cumulative transition distributions, are also in closed form. This property makes CDF-based or quantile-based analyses very convenient. Such examples

---

[1]The Aït-Sahalia (1996b) specification assumes that $\mu_Y(y) = \alpha_{-1}y^{-1} + \alpha_0 + \alpha_1 y + \alpha_2 y^2$ and $\sigma_Y^2(y) = \beta_0 + \beta_1 y + \beta_2 y^{\beta_3}$.

include Value-at-Risk analysis (Jorion 2006), pricing options (Black and Scholes 1973), conditional copula modelling (Patton 2006a,b and 2009), evaluating predictive densities (Diebold et al. 1998), etc. Finally, the monotonicity of the transformations of RDs implies that crucial time series properties of discretely observed RDs such as stationarity, ergodicity and mixing are trivially implied from their BDs. See Doukhan (1994) and Forman and Sørensen (2014) for more detailed discussions.

While RDs potentially have many important advantages, the specifications suggested by Bu et al. (2011) are relatively restrictive compared to the vast literature on nonlinear stochastic modelling. It is quite unlikely that their time-homogeneous structure can be sufficient in describing varied and various empirical dynamics except for only a few very special circumstances. Therefore, useful generalizations of this valuable framework and feasible extensions of existing specifications are extremely important from both theoretical and practical point of views.

The main contribution of this paper is to propose a number of flexible and easy-to-implement extensions to the specification of RDs and examine their empirical performance. Our objective is to generate sufficiently flexible transition densities on the basis of time-homogeneous RDs while maintain sufficient tractability so that classic inferential techniques such as ML estimation can be easily implemented. The concept of conditional time-variation in financial modelling find its root in the pioneering work of Engle (1982) Autoregressive Conditional Heteroskedasticity (ARCH) specification for conditional variances. This insight was then generalized by Hansen (1994) in his general Autoregressive Conditional Density (ACD) framework. Hansen's suggestion is to select a distribution which depends upon a low-dimensional parameter vector, and then allow this parameter vector to vary as a function of the conditional variables. While Hansen's approach assumes that the conditioning set is perfectly adaptive (i.e. observable in the filtration of the process), there is a popular view that the dynamics of economic variables may depend on different states of the world or regimes. This is often referred to as state-dependent dynamic behavior. Depending on whether or not the state of the world at any given point in time is known with certainty in advance, the regime process can be either deterministic or stochastic. The latter case is particularly appealing, since it effectively creates a two factor stochastic process. See Chang et al. (2013) for the latest development in regime-switching stochastic processes.

Although in theory time-variations can be imposed on all elements of the parameter vector of the baseline model, in practice this is not always feasible. On one hand, imposing time variation on too many parameters tends to make reduce the tractability of the model. On the other hand, it reduces the interpretability and economic justification of the econometric model. Thus, in this paper we only allow the transformation parameter to be time-dependent. In other words, we effectively restrict our attention to *RDs with time-varying transformations*. In fact, one potential interpretation of RDs is that the BDs represent the fundamental risk factor and the empirically observed processes are transformed measures of this risk process. In this sense, by allowing the transformations to be time-dependent in our

RDs, this interpretation may be further enhanced and enriched.

Since the philosophy behind our extensions is applicable to all parametric RDs, our exposition will focus on the CIR-CEV RD only due to its parsimony. Another reason of this choice is that the underlying BD (i.e. CIR) has non-Gaussian transition PDF, which to some extent reflects the general need for deviation from the classic Gaussian framework. Moreover, the domain of the CIR-CEV RD can be more conveniently defined on the positive real line than the OU-CEV specification. This property is particularly appealing for financial modelling since the support of many financial variables (e.g. nominal interests) must be positive.

As we will see in Section 3, the transformation function of the CIR-CEV RD depends on a single parameter $\theta$. We therefore propose a total of five distinct time-variation schemes to allow $\theta$ to be time-varying by introducing dynamics of $\theta_t$ to the model. In Model 1, we specify $\theta_t$ as a deterministic function of the first lag of the state variable, i.e. $Y_{t-\Delta}$ where $\Delta$ is the fixed time interval. In Model 2 to 5, we introduce a regime switching mechanism. Specifically, Model 2 is a Self-Exited Threshold (SET) regime switching process where the threshold variable is taken as $Y_{t-\Delta}$. In order to allow regimes-switching to be a continuous process, Model 3 is specified as Logistic Smooth Transition (LST) process. Note that for Model 1 to 3, the time dependence of $\theta_t$ are deterministic. To overcome this limitation, Model 4 and 5 allow the transition of states to follow a two state Markov chain. Model 4 assumes that the transition probability between states are independent of the history of $Y_t$ and hence completely exogenous, whereas Model 5 allows potential nonlinear endogeneity in the transition probabilities. Our benchmark CIR-CEV model is named as Model 0.

As an illustration, we apply our time-varying RD models to UK and US short-rates data. ML was chosen to be our inferential method. For both rates, overwhelming evidence of time-variations in the transformation function was found. Our results suggest that while the naive time-invariant CIR-CEV RD is insufficient in capturing complex dynamics of the data, the proposed time-varying transformation models appear to be very effective in picking up additional variability in the dynamics of the data. In particular, evidence of endogenous regime switching was found in our data, which is consistent with the findings of Chang et al. (2013).

The remainder of the paper is organized as follows. In Section 2, we briefly review the RD framework. Section 3 outlines the details of our baseline model and reviews some of its properties. In Section 4, we present in detail five distinct but representative time-varying transformation schemes. An empirical study based on weekly UK and US short-rates is presented in Section 5. Section 6 concludes.

## 2. Parametric Reducible Diffusions

Bu et al. (2011) proposed to model stochastic processes in finance by parametric RDs. They assume that the dynamics of the state variable $Y_t$ is described by

$$dY_t = \mu_Y(Y_t, \phi)\, dt + \sigma_Y(Y_t, \phi)\, dW_t \tag{2}$$

and crucially (2) belongs to the reducible class defined by Kloeden and Platen (1992). Specifically, the time-homogeneous RD $Y_t$ in (2) is defined as a strictly monotone time-independent transformation of an analytically solvable parametric BD $X_t$, that is

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

where $X_t$ is driven by

$$dX_t = \mu_X(X_t, \omega)\, dt + \sigma_X(X_t, \omega)\, dW_t \tag{3}$$

which depends on parameter vector $\omega$. As such, $V(x, \theta)$ is the transformation function satisfying $\partial V(x, \theta)/\partial x \neq 0$ for all $x$ and then $\theta$ is the transformation parameter vector. Note that the two standard Brownian motions in (2) and (3) are indeed the same Brownian motion, since $V(x, \theta)$ is deterministic and $W_t$ in (3) is the only source of uncertainty. Also, we will usually have $\phi = \omega \cup \theta$ provided that $X_t$ is suitably normalized.

Ito's Lemma determines that

$$\mu_Y(y, \phi) = \mu_X(U(y), \omega) V'(U(y), \theta) + \frac{1}{2}\sigma_X^2(U(y), \omega) V''(U(y), \theta) \tag{4}$$

$$\sigma_Y(y, \phi) = \sigma_X(U(y), \omega) V'(U(y), \theta) \tag{5}$$

where $U(y) = V^{-1}(y)$ (which may be called the *reduction function*) is the unique inverse function of $V(x)$, and $V'(x)$ and $V''(x)$ are the first and second derivatives of $V(x)$ respectively.

The most important task in modelling parametric RDs is the specification of $\mu_Y(y, \phi)$ and $\sigma_Y^2(y, \phi)$ in (2). In this framework, the choice of (3) is typically restricted within the class of parametric diffusions that have closed-form transition PDFs. Hence, $\mu_X(x, \omega)$ and $\sigma_X(x, \omega)$ are often easy to determine. Bu et al. (2011) noted that in theory for a given BD $X_t$, the knowledge of the functional form of any one of $V(x, \theta)$, $\mu_Y(y, \phi)$ or $\sigma_Y^2(y, \phi)$ can lead to unique solutions of the other two functions. This can be easily verified by Ito's Lemma. Thus, there are three different ways of specifying $Y_t$. However, in practice they are not equally feasible. Directly specifying $V(x, \theta)$ may not be straightforward mainly because there is generally a lack of guidance for making such a parametric choice. Starting with a known function $\mu_Y(y, \phi)$ is also infeasible, since one has to solve a 2nd-order Ordinary Differential Equation (ODE) to get $V(x, \theta)$ and $\sigma_Y^2(y, \phi)$. Analytical solutions do not usually exist except for very special cases. Bu et al. (2011) argue that the most analytically tractable

and economically justifiable approach is to start with a desired diffusion function $\sigma_Y^2(y, \phi)$. In this case, one only needs to solve a 1st-order ODE which involves at most a one-dimensional numerical integration. From the economic perspective, the drift term is more related to the long-run behavior of the state variable, whereas the diffusion term is more important in financial applications such as volatility estimation and option pricing which rely mostly on the short-run dynamics of the stochastic process. Moreover, since the drift function is significantly more difficult to estimate than the diffusion term from discrete samples, the preference on the shape of the drift or even the existence of nonlinearity for certain types of data is still debatable (see Choi 2009 for more discussion).

In practice, it is often more convenient to work with the reduction function $U(y)$. Starting from any given $\sigma_Y^2(y, \phi)$, $U(y)$ is simply the solution to the following 1st-order ODE

$$\frac{\partial U(y)}{\partial y} = \frac{\sigma_X(U(y), \omega)}{\sigma_Y(y, \phi)} \tag{6}$$

Solving (6) is comparatively easy and in many cases analytical solutions exist. Once $U(y)$ is identified, the Jacobian of the transformation is in fact given by the right-hand side of (6). Define $\Delta$ as the time interval between neighboring observations and let $f_X(x_t | x_{t-\Delta}; \omega)$ and $f_Y(y_t | y_{t-\Delta}; \phi)$ be the transition PDFs of $X_t$ and $Y_t$, respectively. It follows from the distribution transformation that

$$f_Y(y_t | y_{t-\Delta}; \phi) = |U'(y_t, \theta)| f_X(U(y_t, \theta) | U(y_{t-\Delta}, \theta); \omega)$$

where $U'(y) = \partial U(y)/\partial y$. Let $F_X(x_t | x_{t-\Delta}; \omega)$ and $F_Y(y_t | y_{t-\Delta}; \phi)$ be the corresponding conditional CDFs, we will further have

$$F_Y(y_t | y_{t-\Delta}; \phi) = F_X(U(y_t, \theta) | U(y_{t-\Delta}, \theta); \omega)$$

for $\partial V(x, \theta)/\partial x > 0$ or

$$F_Y(y_t | y_{t-\Delta}; \phi) = 1 - F_X(U(y_t, \theta) | U(y_{t-\Delta}, \theta); \omega)$$

for $\partial V(x, \theta)/\partial x < 0$. Under the assumption that $X_t$ is strictly stationary, so will be $Y_t$. Hence, the above relationships hold for their marginal distributions too.

## 3. CIR-Reducible CEV Diffusion

The CEV diffusion was introduced by Chan et al. (1992) who considered a linear drift term and claimed that it was the best fitting model in their study. The CEV specification was further studied by Aït-Sahalia (1996b) who promoted the use of a nonlinear drift function to provide a better mean reversion effect. The same type of specification was also estimated by Conley et al. (1997) and Gallant and Tauchen (1998). Choi (2009) recently studied a Markov switching CEV diffusion with a nonlinear drift.

### 3.1. Transformation Function and SDE

The CEV diffusion function is given by $\sigma_Y^2\left(y,\phi\right) = \sigma_0^2 y^{2\theta}$ where $\theta \in (0,1) \cup (1,\infty)$. Following the discussion above, for a desired diffusion term and a given BD, the transformation function can be uniquely determined from (6). Bu et al. (2011) suggested that one convenient choice of BD is the square-root (CIR) process which can be written as

$$dX_t = \left(\alpha_2 + \alpha_1 X_t\right) dt + \sigma X_t^{1/2} dW_t \tag{7}$$

where $\alpha_1 < 0$, $\alpha_2 > 0$ and $\sigma > 0$. Using the CIR process as the BD has at least three advantages. Firstly, the CIR process has well known closed-form transition PDF which follows a non-central $\chi^2$ distribution with fractional degrees of freedom. Secondly, when $\sigma_Y^2\left(y,\phi\right) = \sigma_0^2 y^{2\theta}$, the ODE in (6) can be solved analytically and the solution $U\left(y\right)$ is a simple power function. See eq. (8) below. Thirdly, unlike the OU process which has domain on $(-\infty, +\infty)$, the domain of the CIR process can be rigorously restricted to $(0, +\infty)$, which ensures that the required power transformation is strictly monotone and that the domain of $Y_t$ is also on $(0, +\infty)$ as required by many financial variables.

After suitably normalizing the scale of $X_t$ by setting $\sigma_0 = \sigma$, it follows that the CIR-reducible CEV diffusion has the following transformation (reduction) function

$$x = U\left(y,\theta\right) = \left(1/4\right)\left[y^{1-\theta}/\left(1-\theta\right)\right]^2 \tag{8}$$

Since $U'\left(y\right) = y^{1-2\theta}/\left(2-2\theta\right)$, the above transformation is strictly monotone on $(0, +\infty)$ for any given $\theta \in (0,1) \cup (1,\infty)$. The SDE of the CIR-CEV process is then obtained as

$$dY_t = \left\{\left[2\alpha_2\left(1-\theta\right) + \frac{1}{2}\sigma^2\left(2\theta-1\right)\right]Y_t^{2\theta-1} + \frac{\alpha_1 Y_t}{\left(2-2\theta\right)}\right\}dt + \sigma Y_t^\theta dW_t \tag{9}$$

Thus, the parameter vector of the CIR-CEV RD is $\phi = \left(\alpha_1, \alpha_2, \sigma^2, \theta\right)$ in which $\theta$ is the sole transformation parameter.

It is worth mentioning that the drift and diffusion terms of (9) are both nonlinear. In particular, in addition to having a CEV diffusion term, the drift term also exhibits a much stronger pull at high levels of the state variable than the linear drift. Both properties are consistent with empirical findings about the two functions reported in Aït-Sahalia (1996a,b), Conley et al. (1997), Stanton (1997) and many others. Moreover, the SDE in (9) encompasses the AG model with $\theta = 1.5$ and obviously the CIR model itself with $\theta = 0.5$. Clearly, the CIR-CEV process is a more general setup which not only provides the nonlinearity in both terms but also allows for extra degrees of freedom in the data-driven choice of $\theta$.

### 3.2. Statistical Properties

Bu et al. (2011) studied the CIR-CEV process in significant details. In this section, we briefly review some of its most important statistical properties. First of all, the following proposition

regulates the stationarity, boundary behavior, and temporal dependence properties of the CIR-CEV process.

**Proposition 3.1.** *Let $\{Y_t, t \geq 0\}$ be a CIR-CEV process defined in (9). The necessary and sufficient conditions for mean reversion and the unattainability of the boundaries, i.e. 0 and $\infty$, in finite expected time are: (i) $\alpha_1 < 0$ and $4\alpha_2/\sigma^2 > (2\theta - 1)/(\theta - 1)$ if $\theta \in (1, \infty)$; and (ii) $\alpha_1 < 0$ and $4\alpha_2/\sigma^2 > 1/(1 - \theta)$ if $\theta \in (0, 1)$. Under the same conditions, the discretely observed random sequence $\{Y_{i\Delta}, i = 1, ..., n\}$ is strictly stationary and $\beta$-mixing (hence ergodic) satisfying $k^\delta \beta(k) \to 0$ as $k \to +\infty$ for some fixed $\delta > 1$.*

The conditions in (i) and (ii) are valid for arbitrary values of $\theta \in (0, 1) \cup (1, \infty)$. Thus, they are more general than the conditions given by Cox et al. (1985) for the CIR model and Ahn and Gao (1999) for the AG model, which are both special cases of (9) with $\theta = 0.5$ and $1.5$ respectively. Under either (i) or (ii), the CIR-CEV process is strictly positive, mean reverting and stationary. The $\beta$-mixing property ensures that the classical asymptotic theory holds for the ML inference based discretely observed random samples. The proof of this proposition follows straightforwardly from the regularity conditions set out in Aït-Sahalia (1996b). See Bu et al. (2011) for more details.

Secondly, the next proposition summarizes the distributional properties of discretely observed random sequence of the CIR-CEV process.

**Proposition 3.2.** *Under condition (i) or (ii), the transition PDF of the CIR-CEV process is unique and can be expressed in closed form as*

$$f_Y(y_t|y_{t-\Delta};\phi) = \frac{1}{2} \frac{y_t^{1-2\theta}}{|1-\theta|} ce^{-u-v} \left(\frac{v}{u}\right)^{q/2} I_q \left[2(uv)^{1/2}\right] \tag{10}$$

*where $c = 2\alpha_1 / \left[\sigma^2 \left(e^{\alpha_1\Delta} - 1\right)\right]$, $u = \left(ce^{\alpha_1\Delta}/4\right) \left[y_{t-\Delta}^{1-\theta}/(1-\theta)\right]^2$, $v = (c/4) \left[y_t^{1-\theta}/(1-\theta)\right]^2$, $q = 2\alpha_2/\sigma^2 - 1$ and $I_q(\cdot)$ is the modified Bessel function of the first kind of order $q$. The corresponding closed-form conditional CDF is given by*

$$F_Y(y_t|y_{t-\Delta};\phi) = \begin{cases} D(2cx_t; 2q+2, 2u) & \text{for } \theta \in (0, 1) \\ 1 - D(2cx_t; 2q+2, 2u) & \text{for } \theta \in (1, \infty) \end{cases} \tag{11}$$

*where $x_t$ is given in (8) and $D(\cdot; 2q+2, 2u)$ is the non-central $\chi^2$ distribution with $2q+2$ degrees of freedom and non-centrality parameter $2u$. Meanwhile, the marginal distribution of the CIR-CEV process exists and is given by*

$$\pi(y;\phi) = \frac{1}{2} \frac{y^{1-2\theta}}{|1-\theta|} g\left(x; \frac{2\alpha_2}{\sigma^2}, -\frac{2\alpha_1}{\sigma^2}\right) \tag{12}$$

10

where $x$ is given in (8) and $g\left(\cdot; 2\alpha_2/\sigma^2, -2\alpha_1/\sigma^2\right)$ is the PDF of the Gamma distribution with shape parameter $2\alpha_2/\sigma^2$ and scale parameter $-2\alpha_1/\sigma^2$. The corresponding closed-form marginal CDF is given by

$$\Pi_Y\left(y; \phi\right) = \begin{cases} G\left(x; \frac{2\alpha_2}{\sigma^2}, -\frac{2\alpha_1}{\sigma^2}\right) & \text{for } \theta \in (0,1) \\ 1 - G\left(x; \frac{2\alpha_2}{\sigma^2}, -\frac{2\alpha_1}{\sigma^2}\right) & \text{for } \theta \in (1,\infty) \end{cases} \tag{13}$$

where $G\left(\cdot; 2\alpha_2/\sigma^2, -2\alpha_1/\sigma^2\right)$ is the Gamma CDF.

The closed-form transition PDF in (10) allows the users to implement exact ML inference without any computational burden. This is in contrast to the majority of parametric diffusions in the literature (cf. Durham and Gallant 2002). Moreover, the closed-form conditional CDF in (11) is a convenient tool for various statistical and financial applications which rely essentially on conditional quantiles (cf. Jorion 2006, Patton 2009, Diebold et al. 1998). Generally speaking, all parametric RDs have at least these advantages by their construction.

Finally, the explicit close-form expressions the conditional and unconditional moment functions of discretely observed random sequence of the CIR-CEV process are provided in the next proposition.

**Proposition 3.3.** *Under condition (i) or (ii), the conditional moment functions of the CIR-CEV process are given by*

$$E\left(y_t^m | y_{t-\Delta}; \phi\right)$$
$$= \left[2\left|1-\theta\right|\right]^{\frac{m}{1-\theta}} c^{-\frac{m}{2(1-\theta)}} e^{-u} \frac{\Gamma\left(q + \frac{m}{2(1-\theta)} + 1\right)}{\Gamma\left(1+q\right)} {}_1F_1\left(q + \frac{m}{2\left(1-\theta\right)} + 1, 1+q, u\right) \tag{14}$$

*for $m > 0$, where ${}_1F_1\left(\cdot, \cdot, \cdot\right)$ is the confluent hypergeometric function[2] defined as*

$${}_1F_1\left(a, b, y\right) = \sum_{j=0}^{\infty} \frac{(a)_j\, y^j}{(b)_j\, j!} = \frac{\Gamma\left(b\right)}{\Gamma\left(b-a\right)\Gamma\left(a\right)} \int_0^1 e^{yz} z^{a-1}\left(1-z\right)^{b-a-1} dz$$

*with $\Gamma(\cdot)$ being the Gamma function. The corresponding unconditional moment functions are given by*

$$E\left[y^m; \phi\right] = \left[2\left|1-\theta\right|\right]^{\frac{m}{1-\theta}} \left(-\frac{2\alpha_1}{\sigma^2}\right)^{-\frac{m}{2(1-\theta)}} \frac{\Gamma\left(\frac{2\alpha_2}{\sigma^2} + \frac{m}{2(1-\theta)}\right)}{\Gamma\left(\frac{2\alpha_2}{\sigma^2}\right)} \tag{15}$$

These explicit moment functions are potentially very useful for implementing moment-based estimations and goodness-of-fit tests. In particular, $m$ in (14) and (15) is allowed to be any positive real numbers. Thus, (14) and (15) in fact permit the evaluation of the fractional moments of the CIR-CEV process. A sketch of the proofs of (10) to (15) can be found in Bu et al. (2011).

---

[2]Abadir (1999) provided a detailed account of the hypergeometric functions and their applications in economics. Also, see Abadir and Rockinger (2003) and Bu and Hadri (2007) for applications in finance.

# 4. Time-Varying Transformations

The RD framework of Bu et al. (2011) was developed under the assumption that time-invariant transformations are imposed on time-homogeneous BDs to form time-homogeneous RDs. Although this modelling strategy is original, the maintained assumption of time-homogeneity may prove to be restrictive from the practical point of view[3]. In their empirical application of OU-CEV and CIR-CEV models, they found evidence of time-variation in their transformations. There are two ways to make RDs (or more generally transformation models) more flexible. One way is to use more flexible underlying processes. However, in the context of diffusions, the choices of models with closed-form solutions are quite limited, which is in fact the motivation behind RD modelling in the first place. Therefore, a more feasible and practical solution is to introduce suitable time-variation dynamics to the transformations.

Focusing on the baseline CIR-CEV model, we can see that the transformation depends on a single transformation parameter $\theta$. Thus, a natural strategy is to consider various feasible and practical nonlinear time-varying mechanisms for describing the evolution of $\theta$ over time. We consider our approach feasible and practical in terms of both econometric tractability and empirical flexibility. That is, we aim to create models that can account for empirical dynamics as flexibly as possible, but at the same time we prefer to enable standard inferential methods such the ML to be implemented without too much difficulties following the introduction of additional nonlinear features.

Generally speaking, time-varying parameter models can be broadly divided into two classes. The first class assume that the value of the varying parameter at time $t$ is non-stochastically determined by values of variables observed up to time $t$. In contrast, the other class assume that the time-varying parameter values cannot be observed but are determined by a second unobservable stochastic process. This implies that one can never be certain about the value of the time-varying parameter at any particular point in time, but can only assign probabilities to the occurrence of different values. In this paper, we consider a total of five widely accepted time-varying specifications representing both classes of time dependence in order to reflect a variety of economically and statistically justifiable situations.

## 4.1. Autoregressive Conditional Transformation (Model 1)

The ARCH specification of Engle (1982) was one of the earliest innovations in finance for considering conditional time dependence. This concept of autoregressive conditioning was then generalized by Hansen (1994) in modelling conditional densities. The suggestion is to model conditional densities (instead of just the mean and the variance) with a small number of parameters and then model these parameters as functions of the conditioning information.

In the current context, to account for potential time-variation in transformation parameter $\theta$, we assume that $\theta_t$ follows the following Autoregressive Conditional Density (ACD)

---

[3]See Choi (2013) for some empirical evidence of time-inhomogeneity in diffusion modelling.

type dynamics

$$\theta_t = \widetilde{\Lambda}\left(\omega + \sum_{i=1}^{p} \beta_i Y_{t-\Delta}^i\right) \tag{16}$$

where the link function $\widetilde{\Lambda}(z) \equiv (1 + e^{-z})$ is the reciprocal of the logistic function to restrict $\theta_t$ on the domain of $(1, \infty)$ at all times. We can also let the link function to be the logistic function $\Lambda(z) \equiv (1 + e^{-z})^{-1}$ to keep $\theta_t$ on $(0, 1)$.

As can be seen from (16), we restrict our forcing variable at time $t$ to be the observation at time $t - \Delta$, i.e. $Y_{t-\Delta}$. This choice has the advantage of maintaining the first-order Markovian property of the resulting model, which is one of the fundamental assumptions of diffusion modelling. In addition, to account for the possibility of nonlinear dependence on $Y_{t-\Delta}$, we consider a polynomial function of $Y_{t-\Delta}$ and expect that the nonlinear dependence of $\theta_t$ on $Y_{t-\Delta}^i$, if any, can be picked up by the significance of coefficient $\beta_i$. This is analogous to the reasoning behind the RESET test (cf. Ramsey 1969). Moreover, our polynomial design and the subsequent ML estimation is also analogous to the sieve estimator in semiparametric statistics (cf. Ai and Chen 2003). Thus, to some extent Model 1 has some semiparametric interpretation. In practice, the choice of $p$ should be decided by some model selection criteria.

### 4.2. Self-Exited Threshold Transformation (Model 2)

Regime switching models have been used extensively in economic modelling. In most studies, two regimes designated as high (H) and low (L) states of an economy are introduced with a state process making one of the regimes take place in each period. A most prominent member of regime switching models is the threshold model, initially proposed by Tong (1978) and Tong and Lim (1980), and discussed extensively in Tong (1990), which assumes that the regime that occurs at time $t$ can be determined by an observable variable relative to some threshold value. In univariate time series modelling, a special case arises when the threshold variable is taken to be a lagged value of the time series itself, i.e. $Y_{t-d\Delta}$ for a certain integer $d > 0$. Since in this case the regime is determined by the time series itself, the resulting model is called a Self-Excited Threshold (SET) model.

In the context of RDs with conditional transformations, we assume that the value of $\theta_t$ changes depending on whether $Y_{t-\Delta}$ $(d = 1)$ is above an unknown threshold value $\tau_{SET}$ or not. Thus, the dynamics of $\theta_t$ can be written as

$$\theta_t = \begin{cases} \theta_L \text{ if } Y_{t-\Delta} \leq \tau_{SET} \\ \theta_H \text{ if } Y_{t-\Delta} > \tau_{SET} \end{cases} \tag{17}$$

An alternative way to write the SET model in (17) is

$$\theta_t = \theta_L \left(1 - I\left[Y_{t-\Delta} > \tau_{SET}\right]\right) + \theta_H I\left[Y_{t-\Delta} > \tau_{SET}\right]$$

where $I[A]$ is an indicator function with $I[A] = 1$ if the event $A$ occurs and $I[A] = 0$ otherwise. See Tong (1990) for more detailed analyses of this specification.

13

### 4.3. Smooth Transition Transformation (Model 3)

The SET model assumes that the border between the two regimes is given by a specific value of the threshold variable $Y_{t-\Delta}$. A more gradual transition between different regimes can be obtained by replacing the indicator function by a continuous transition function which changes smoothly from 0 to 1 as $Y_{t-\Delta}$ increases. Unsurprisingly, a popular choice for the transition function is the logistic function with

$$\Lambda(Y_{t-\Delta}; \lambda, \tau_{LST}) = \frac{1}{1 + \exp\left[-\lambda\left(Y_{t-\Delta} - \tau_{LST}\right)\right]} \tag{18}$$

and the dynamics of $\theta_t$ is written as

$$\theta_t = \theta_L\left(1 - \Lambda(Y_{t-\Delta}; \lambda, \tau_{LST})\right) + \theta_H \Lambda(Y_{t-\Delta}; \lambda, \tau_{LST}) \tag{19}$$

The resulting model is called a Logistic Smooth Transition (LST) model.

The parameter $\tau_{LST}$ in (18) can be interpreted as the threshold between the two regimes corresponding to $\Lambda(-\infty; \lambda, \tau_{LST}) = 0$ and $\Lambda(+\infty; \lambda, \tau_{LST}) = 1$, in the sense that the logistic function changes monotonically from 0 to 1 as $Y_{t-\Delta}$ increases, while $\Lambda(\tau_{LST}; \lambda, \tau_{LST}) = 0.5$. The parameter $\lambda$ determines the smoothness of the change in the value of the logistic function, and thus the transition from one regime to the other. It is easily verified that as $\lambda$ becomes very large, the change of $\Lambda(Y_{t-\Delta}; \lambda, \tau_{LST})$ from 0 to 1 becomes almost instantaneous at $Y_{t-\Delta} = \tau_{LST}$ and, consequently, the logistic function $\Lambda(Y_{t-\Delta}; \lambda, \tau_{LST})$ approaches the indicator function $I\left[Y_{t-\Delta} > \tau_{LST}\right]$. Hence the SET model in (17) can be approximated by the LST model in (19). When $\lambda \to 0$, the logistic function becomes equal to a constant (equal to 0.5) and when $\lambda = 0$, the LST model reduces to the time-invariant baseline model. See Teräsvirta (1994) and others for more details of this specification.

### 4.4. Exogenous Regime Switching Transformation (Model 4)

The time-varying specifications of $\theta_t$ discussed above are all deterministic with respect $t$. However, 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. (2013) and so on. A most recent study of regime switching diffusions was by Choi (2009).

Typically, in the context of diffusions, a continuous-time two state Markov chain with the conservative infinitesimal matrix

$$Q = \begin{pmatrix} q_{LL} & q_{HL} \\ q_{LH} & q_{HH} \end{pmatrix} = \begin{pmatrix} -q_{LH} & q_{HL} \\ q_{LH} & -q_{HL} \end{pmatrix}$$

is assumed to govern the switching between two regimes. The intensity parameter $q_{ij}$ is the rate of the probability at which the process switches from the state $i$ to the state $j$ as time

goes to zero and $q_{ij} > 0$ for $i \neq j$. The corresponding transition matrix is then

$$P^\Delta = \frac{1}{q_{LH} + q_{HL}} \begin{pmatrix} q_{HL} + q_{LH}e^{-\Delta(q_{LH}+q_{HL})} & q_{HL}\left(1 - e^{-\Delta(q_{LH}+q_{HL})}\right) \\ q_{LH}\left(1 - e^{-\Delta(q_{LH}+q_{HL})}\right) & q_{LH} + q_{HL}e^{-\Delta(q_{LH}+q_{HL})} \end{pmatrix}$$

where $\Delta$ is the time interval between two neighboring observations. Hence, the transition probabilities depend not only on the intensity but also on $\Delta$. Moreover, $\pi_L = q_{HL}/(q_{LH} + q_{HL})$ and $\pi_H = q_{LH}/(q_{LH} + q_{HL})$ are the unconditional probabilities that the above Markov chain will be in state $L$ and $H$, respectively at any time.

If $\Delta$ is small enough, we can assume approximately that at most one regime shift can occur in the duration of $\Delta$. Thus, the regime index $s_t$ is assumed to follow a discrete-time Markov chain with two states. As such, we can reparameterize $P^\Delta$ as

$$P = \begin{pmatrix} p_{LL} & p_{HL} \\ p_{LH} & p_{HH} \end{pmatrix} \tag{20}$$

and the transition probabilities, $p_{ij} = P(s_t = j | s_{t-\Delta} = i)$, $i, j = L, H$ are of greater interest than the intensity parameters in understanding the dynamics of stochastic processes. Representing stationary distribution of the Markov chain in terms of $p_{LL}$ and $p_{HH}$, we have

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

which is, not surprisingly, the same as the unconditional probabilities of a discrete-time Markov chain with transition matrix (20). Consequently, the transition probabilities $p_{LL}$ and $p_{HH}$ and the parameters of regime-dependent RDs can be estimated in exactly the same way as discrete-time Markov switching models using a Hamilton (1989) type filtering algorithm[4]. Meanwhile, in practice we specify the transition probabilities as $p_{LL} = \Lambda(c_L)$ and $p_{HH} = \Lambda(c_H)$ to ensure that $p_{LL}, p_{HH} \in (0, 1)$.

### 4.5. Endogenous Regime Switching Transformation (Model 5)

The above standard regime switching model assumes that the transition of the underlying finite state Markov chain is independent of the evolutionary path of the observed state variable $Y_t$. However, studies including Diebold et al. (1994), Kim et al. (2008), Choi (2009), Chang et al. (2013) among others all reported evidence of endogeneity in regime changes that we observe widely and frequently for many practical applications. A convenient way to allow endogenous regime changes in regime switching models is to model the transition probabilities as functions of the lagged values of the state variable $Y_t$ itself. In order to maintain the Markovian property of the resulting model and allow for sufficient nonlinearity

---

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

in the dependence, we adopt a similar specification as in Model 1 for the two endogenous transition probabilities as

$$p_{LL}\left(Y_{t-\Delta}\right) \quad = \quad P\left(s_t = L|s_{t-\Delta} = L, Y_{t-\Delta}\right) = \Lambda\left(c_L + \sum_{i=1}^{p} d_{L,i} Y_{t-\Delta}^i\right) \qquad (21)$$

$$p_{HH}\left(Y_{t-\Delta}\right) \quad = \quad P\left(s_t = H|s_{t-\Delta} = H, Y_{t-\Delta}\right) = \Lambda\left(c_H + \sum_{i=1}^{p} d_{H,i} Y_{t-\Delta}^i\right) \qquad (22)$$

where $\Lambda$ is the logistic function[5]. Clearly, if all coefficients of the powers of $Y_{t-\Delta}$ are jointly zero, the transition probabilities become constant and the time-varying transition matrix is reduced to the time-invariant transition matrix (20).

It is worth mentioning that our specifications of the endogenous transition probabilities in (21) and (22) are related to the recent approach by Chang et al. (2013). They assume that the switching between regimes is determined by whether a latent process is above or below some unknown threshold value. Crucially, they assume that the innovation term of the latent process and the observed state variable are correlated. As a consequence, the transition probabilities of their regime switching model become endogenous, that is, dependent on the lag(s) of the state variable. Clearly, our specifications in (21) and (22) produce similar effects and are therefore consistent with the logic behind their methodology. However, in contrast to their implicit approach based on a latent process and an unknown threshold, we choose to directly and explicitly specify $p_{LL}$ and $p_{HH}$ as flexible functions of $Y_{t-\Delta}$ to introduce similar endogeneity. Nevertheless, our approach clearly cannot give any intuitive description of the cause of endogeneity as their model does.

For both Model 4 and 5, we define $\theta_L$ as the low-value transformation parameter (corresponding to the low state) and $\theta_H$ as the higher value transformation parameter (corresponding to the high state). It is important to emphasize that this definition of $\theta_L$ and $\theta_H$ is different from the one in Model 2 and 3 where, according to (17) and (18), $\theta_L$ is defined as the transformation parameter corresponding to $Y_{t-\Delta}$ below or equal to the threshold and $\theta_H$ is defined corresponding to $Y_{t-\Delta}$ above the threshold.

## 5. Application to UK and US Short-Term Interest Rates

Short-term interest rates are often modeled as continuous-time diffusion processes. Perhaps more diffusion specifications have been proposed for modelling short-rates than any other financial variables. However, most earlier models are based on time-homogeneous parametric single-factor diffusions, which prove to be insufficient for describing complex short-rate

---

[5]The choice of link functions here is not unique. Other options include the standard normal CDF function and others. However, results are usually quite simiar which is true in our applications too.

dynamics. In this section, we examine the empirical performance of RDs with time-varying transformations proposed in this paper.

## 5.1. The Data

We measure the UK and the US short rates by the 1-Month London Interbank Offered Rate (LIBOR) in British Sterling and the 1-Month Eurodollar Rate (EDR) respectively, both at the weekly frequency. The LIBORs are obtained from BBA (British Banking Association) database and the EDRs are collected from the H.15 release of the Federal Reserve website[6].

Table 1 provides some summary statistics of the data. The LIBOR data are available from January 1986 to December 2007 with 1148 weekly observations. The EDR data start from January 1971 to December 2007 yielding 1930 weekly observations. The sample means of the two series suggest that the LIBOR is on average higher than the EDR, whereas the standard deviations indicate that the EDR is more volatile than the LIBOR. The departure from normality is confirmed by the significance of Jarque-Bera normality tests on the marginal distributions. This suggests that any models, continuous or discrete, that imply Gaussian marginal distributions will not be entirely appropriate. While both rates show positive skewness, we note that the LIBOR is platykurtic whereas the EDR is leptokurtic.

The time series plots of the two series and their first differences are provided in Figure 1. Neither of the two series show discernible trend over its sample period. For the EDR, due to the shift in monetary policy, the 1980 to 1982 years are characterized by substantially higher levels than the rest of the sample period. Associated with high levels of interest rates are high levels of volatility. In fact, this observation has been a main reason that motivates the CEV diffusion specification by many authors (cf. Chan et al. 1992, Conley et al. 1997). The LIBOR is highest around year 1990 as a result of the monetary policy followed at the time. Both rates reached their lowest levels during the years 2004 and 2005. Generally speaking, some graphical evidence exists which suggests that short-rates behave quite differently in different time periods. It is particularly clear from the lower panel of Figure 1 that the volatilities are time-dependent. We conjecture that time-variations with possibly regime changes are likely to be an important feature of our data.

## 5.2. ML Estimation

One of the most prominent advantages of RDs is their closed-form transition PDFs. Thus, ML is chosen to be our preferred estimation method for its efficiency gain and classic inferential procedure. The ML estimation for Model 0 and Model 1 are straightforward. For

---

[6]We use the same dataset used by Bu et al. (2011) so that their estimation results can serve as a benchmark.

Model 2, we search for the optimal threshold value $\tau_{SET}$ across the full range of the sample[7]. Specifically, we take each distinct observation in the sample one at a time as a candidate threshold and use the ML procedure to obtain the corresponding log-likelihood (LL) of the model. We repeat this procedure using every distinct observation as the threshold. The estimation results corresponding to the highest LL are then reported[8]. For Model 3, the parameter space is assumed to be continuous. However, in order to find the global maximum, we adopt the following procedure to search for the optimal $\tau_{LST}$. We first use a similar search procedure as in Model 2 to locate the candidate threshold which yields the highest LL for the restricted version of Model 3 (i.e. Model 3 with known threshold). We then use this candidate threshold and the estimates of other parameters as starting values to estimate the unrestricted version of Model 3. The results obtained in this way are then reported. Model 4 and 5 are hidden Markov regime switching models. Therefore, their transition PDFs, i.e. $f_Y(y_t|y_{t-\Delta})$, are not immediately available since it depends on the unobservable state. In this paper, we use the standard filtering algorithm of Hamilton (1989) to obtain $f_Y(y_t|y_{t-\Delta})$ and implement ML estimation.

It is important to stress that since the focus of this paper is on time-varying transformations we assume that the underlying BD (i.e. the CIR process) is time-homogeneous. Of course, one can in theory allow additional parameters to be time-dependent to achieve even better flexibility, but a price must be paid for additional difficulty in numerical stability and interpretation of the model. On the other hand, depending on whether $\theta_t \in (1, \infty)$ or $\theta_t \in (0, 1)$, fairly general scenarios may be considered in the context of the models discussed in Section 4. However, for model simplicity and space economy, we assume $\theta_t \in (1, \infty)$ for all $t$ in subsequent analysis. This simplifying assumption is consistent with the empirical findings of Bu et al. (2011).

### 5.3. Results

Estimation results for the two short-rate series (both measured as percentages, i.e. multiplied by 100) are reported in Table 2 and 3, respectively. A total of six models (including Model 0 which is our baseline model) were estimated. For each model, we report ML estimates of model parameters and their standard errors (whenever applicable) obtained by inverting the numerical Hessian. We also report the maximized LL and the corresponding AIC and BIC values. In order to test the significance of time-variation in the transformation, the Likelihood Ratio (LR) statistics and the corresponding $p$-value for each time-varying specification against the baseline model are also reported in the tables. However, it is important to stress that not all of these $p$-values are strictly valid since under the null of Model 0 the

---

[7]Usually it would be sufficient to search the middle 70% of the data as threshold at either end 15% is not too appealing from either economic or econometric point of views.

[8]This procedure can be regarded as a special grid search procedure in the sense that the grid points we use simply consist of all distinct observations in the sample.

parameters of some time-varying specifications are unidentified. Nevertheless, in such cases the $p$-values are only used as crude indications of comparative model performances. More details are explained in the sequel.

Results from the baseline Model 0 are identical to those of Bu et al. (2011). The constant transformation parameter $\theta$ is significantly different from either 0.5 or 1.5. Thus, the nested CIR and AG models are strongly rejected.

Model 1 has the ACD type time-varying specification. The polynomial type dependence on the first lag $Y_{t-\Delta}$ inside the link function captures potential nonlinearity. We report the results for $p = 3$ for both data series. For the LIBOR, according to the $t$-ratios we find that only the coefficients of $Y_{t-\Delta}^2$ and $Y_{t-\Delta}^3$ are significant. For the EDR, the coefficients of $Y_{t-\Delta}$, $Y_{t-\Delta}^2$ and $Y_{t-\Delta}^3$ are significant. This provides clear evidence of significant time-variation in the transformation parameter. However, including higher powers of $Y_{t-\Delta}$ did not significanly improve the goodness-of-fit. In fact, $Y_{t-\Delta}^4$ turned out to be insignificant for both data series. The LL, AIC and BIC all improved significantly from the baseline model. Since Model 1 nests Model 0, we calculated the LR statistic to jointly test the significance of this particular time-varying specification. The resulting LR statistics are 95.46 for LIBOR and 37.64 for EDR, respectively. Compared to the $\chi^2$ distribution with 4 degrees of freedom, the $p$-values are zero up to at least 4 decimal points, suggesting strong rejection of the baseline model. To graphically demonstrate the time-varying effects, we plot the fitted values of $\theta_t$ against the constant $\theta$ estimated from Model 0 as well as its 95% confidence bands in Figure 2. The significant proportion of $\theta_t$ outside the confidence bands are strong evidence of the time-variation of $\theta_t$.

Model 2-5 all assume that $\theta_t$ can take potentially two distinct values corresponding to the two states of the world. Such an assumption is economically intuitive since economies are often viewed to have high or low status. Model 2 is a simple SET model of order 1. As discussed above, we searched the whole sample space for the value of the threshold $\tau_{SET}$ to locate the maximum LL. The optimal $\tau_{SET}$ is found to be 12.625 for the LIBOR and 11.610 for the EDR. For the LIBOR, the transformation parameter is higher ($\theta_L = 1.649$) when $Y_{t-\Delta}$ is below the threshold and lower ($\theta_H = 1.394$) when $Y_{t-\Delta}$ is above the threshold. In contrast, the opposite situation is observed for the EDR with $\theta_L = 1.123$ and $\theta_H = 1.268$. The LL, AIC and BIC all suggest significant improvements upon the baseline model for both rates. However, the LR test cannot be used here since under the null of no threshold effect, $\tau_{SET}$ and the two transformation parameters $\theta_H$ and $\theta_L$ are unidentified and the resulting LR statistic will not have the usual $\chi^2$ distribution. Nevertheless, if we crudely resort to the conventional LR test, Model 0 would be rejected in favor of Model 2 even at 0.01% significance level, indicating some though not strictly formal evidence in favor of the threshold type regime switching effects for both LIBOR and EDR. Coincidentally, the goodness-of-fit of Model 1 and 2 in terms of LL, AIC and BIC are surprisingly close, despite the fact that the time-varying mechanisms of the two models are quite different.

Model 3 is the threshold model with smooth transition between $\theta_H$ and $\theta_L$. Related

19

literature suggests that ST models are often difficult to estimate. In particular, the accuracy in estimating $\lambda$ is usually poor. Unfortunately, this is true in our application too. As expected, the estimates of $\lambda$ are extremely large but insignificant for both series. In the mean time, other parameter estimates are fairly close to those of Model 2, the SET model. As discussed in Section 4, the ST model can approximate the SET model with very large values of $\lambda$. For this reason, the estimation results of Model 3 have significant implications. Clearly, they are indicative that for both rates the transition between $\theta_H$ and $\theta_L$ is unlikely to be smooth. More specifically, the two series are both in favor of the type of transition as specified in Model 2. The fairly similar goodness-of-fit measures of Model 2 and 3 in terms of LL, AIC, and BIC are further confirmation of this observation. A similar informal LR statistic comparing Model 3 and 0 and its $p$-value were also calculated. As expected, they are practically the same as those of Model 2. Thus, unsurprisingly Model 3 also offers strong evidence against the baseline model.

Model 4 and 5 differ from Model 0-3 in that they represent stochastic transformations and thus they are effectively two factor random processes. In Model 4, the transition probabilities between high and low regimes ($p_{HH}$ and $p_{LL}$) are constant. In other words, the switching between regimes is exogenous. However, it is worth stressing that the regime switching itself still determines that $\theta_t$ is time-varying. Estimation results from the two rates suggest that the introduction of random regime switching (stochastic time-variation) substantially improved the goodness-of-fit to the data in a way that is substantially better than any of the deterministic type time-varying transformation models. For both series, the LL, AIC and BIC measures almost all doubled compared to Model 0 and significantly better than Model 1, 2 and 3. Similar to Model 2 and 3, it is also impossible to use the usual LR statistics to test our regime switching model against the time-homogeneous case. A similar reason is that the parameters related to the second state of the process are not identified under the null of no regime-switching. There are some researchers who addressed this problem, but generally testing for multiple regimes is not an easy task (cf. Davis 1987 and Hansen 1992, 1996). However, a similar informal test (as we did for Model 2 and 3) based on the usual LR statistic and $\chi^2$ distribution suggests an even stronger rejection of Model 0 in favor of Model 4 at less than 0.01% significance level. In any event, the substantially improved AIC and BIC are more convincing indications.

For the LIBOR, the estimated transition probabilities $p_{LL} = 0.877$ and $p_{HH} = 0.835$ suggest that the average number of periods that $\theta_t$ stays in the low value is approximately 8 weeks and 6 weeks in high value. This is slightly in contrast to the EDR with $p_{LL} = 0.967$ and $p_{HH} = 0.933$ and averages of 30 and 15 weeks respectively. Hamilton (1989)'s algorithm also allows us to infer the probability of the state at each time period conditional on the full sample. Such inferred probabilities are also known as the smoothed probabilities. In order to have an indication about the likelihood of $\theta_t$ in either states at each time step, we plot in Figure 3 the time series of smoothed probabilities against the time series of $Y_t$ itself and its first difference. As far as the EDR is concerned, we find that our inferred states coincided

to a reasonable degree with economic events happened in the sample period depicted in Choi (2009)[9]. This suggests that our regime switching time-varying transformation model can indeed reasonably pick up important features of the data. Moreover, compared to the differenced series, distinct volatility levels appear to be characteristic features of different states. Similar features are also observed for the LIBOR.

Although the exogenous regime switching Model 4 performs distinctively well compared to all preceding cases, the assumption of exogenous regime switching contradicts the evidence of endogeneity in regime changes that we observe widely and frequently for many practical applications (cf. Chang et al. 2013)[10]. Model 5 examines the possibility of such endogeneity by specifying $p_{HH}$ and $p_{LL}$ as nonlinear functions of $Y_{t-\Delta}$. For the same argument as for Model 1, we include powers of $Y_{t-\Delta}$ up to 3rd order and use the same algorithm to obtain ML estimation results. For both rates, the goodness-of-fit of Model 5 improved upon Model 4 in terms of LL and AIC, but not BIC. This is not surprising because BIC imposes a heavier penalty on increased number of parameters than AIC and Model 5 has 8 additional parameters relative to Model 4. Both LR statistics calculated against Model 0 are unsurprisingly higher. As far as individual coefficients are concerned, no clear significance was found by examining the $t$-ratios alone. However, since the regressors are highly correlated, looking at $t$-ratios individually may not be totally appropriate. Hence, we perform LR tests to examine all coefficients jointly and test against Model 4. Note that this is actually a valid LR test since Model 4 is nested in Model 5. The resulting LR statistics are 36.97 for LIBOR and 21.06 for EDR respectively, which are both highly significant at 1% significance level. Hence, despite the seemingly insignificant coefficients, our LR tests turn out to be very significant for both LIBOR and EDR. This can be regarded as evidence of nonlinear endogeneity in the regime switching probabilities. To demonstrate this endogeneity graphically, we plot in Figure 4 the filtered endogenous transition probabilities $p_{HH}(Y_{t-\Delta})$ and $p_{LL}(Y_{t-\Delta})$ versus the filtered constant transition probabilities $p_{HH}$ and $p_{LL}$ of Model 4 and the corresponding 95% confidence bands. The presence of time-variation in these filtered transition probabilities is quite obvious. Above all, Model 4 and 5 both offered very strong evidence of regime switching type time-variation in $\theta_t$ against other types of specifications. Moreover, evidence coming out of Model 5 also suggests that such regime switches are likely to be endogenously driven. This result is consistent with recent findings made by Chang et al. (2013) in their application to US GDP growth rates and NYSE/AMEX index returns.

---

[9]Choi (2009) gave a more detailed account of such economic events. For space economy, we do not repeat here.

[10]Chang et al. (2013) provides a more specific and economically justifiable alternative framework for endogenous regime switching dynamics.

## 6. Conclusion

In this paper, we proposed the general idea of using RDs with nonlinear time-varying transformations for modelling financial and economic variables. Just as most pure time series models, the economic intuition behind our model is not immediately evident. However, to a limited extent one may regard the BD as a fundamental economic factor which drives the evolution of observable variables and the transformation can be viewed as a nonlinear measure of this factor. Moreover, the measure itself may be influenced by changing economic environment and therefore the transformation itself can also evolve over time. We took a feasible and practical approach in the design of our RDs, so that they are potentially very flexible but analytically very tractable. Both are appealing properties to empirical users. Using the CIR-CEV process as our baseline model, we proposed five distinct econometric specifications to accommodate a variety of economically justifiable nonlinear time-varying dynamics for the transformation function. These specifications represent both deterministic and stochastic time dependences. Without violating the Markovian property of diffusion models, we only allowed the dependence of time-variation on the first lag of the observed series. Whenever is feasible and convenient, such a lag-dependence is specified in a flexible form which is to some extent analogous to a nonparametric design. Such a consideration further enhances the flexibility and captures potential nonlinearity. The application of our specifications to UK and US short-term interest rate data revealed strong evidence of time-variations in the transformation of our RD. Consistent with recent findings by Chang et al. (2013), the endogenous regime switching model appears to be our best performing specification. Above all, our study suggests that in the context of RDs, time-dependent transformations are either a phenomenon or a statistical tool that may be extremely important for empirical studies.

## 7. Acknowledgement

## References

Abadir, K. M. (1999). An Introduction to Hypergeometric Functions for Economists. Econometric Reviews, 18(3), 287-330.

Abadir, K. M., and Rockinger, M. (2003). Density Functionals, with An Option-pricing Application. Econometric Theory, 19, 778-811.

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

Ai, C., and Chen, X. (2003). Efficient Estimation of Models with Conditional Moment Restrictions Containing Unknown Functions. Econometrica, 71(6), 1795-1843.

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.

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

Brennan, M. and Schwartz, E. (1979). A Continuous Time Approach to the Pricing of Bonds. Journal of Banking and Finance, 3, 133-155.

Bu, R., Giet, L., Hadri, K., and Lubrano, M., (2011). Modelling Multivariate Interest Rates using Time-Varying Copulas and Reducible Non-Linear Stochastic Differential Equations. Journal of Financial Econometrics, 9(1), 198-236.

Bu, R., and Hadri, K. (2007). Estimating Option Implied Risk-neutral Densities using Spline and Hypergeometric Functions. Econometrics Journal, 10, 216-244.

Bu, R., Hadri, K. and Kristensen, D. (2014). Copula-Based Semiparametric Diffusion Models: Identification and Estimation. Working Paper.

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. (2013). Regime Switching Model with Endogenous Autoregressive Latent Factor. Working Paper.

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

Choi, S. (2013). Closed-Form Likelihood Expansions for Multivariate Time-Inhomogeneous Diffusions. Journal of Econometrics, 174, 45-65.

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.

Constantinides, G. M., and Ingersoll, J. (1984). Optimal Bond Trading with Personal Taxes. Journal of Financial Economics, 13, 299-335.

Constantinides, G. M. (1992). A Theory of the Nominal Term Structure of Interest Rates. Review of Financial Studies 5, 531-522.

Courtadon, G. (1982). The Pricing of Options on Default-free Bonds. Journal of Financial and Quantitative Analysis, 17, 75-100.

Cox, J. (1975). Notes on Option Pricing I: Constant Elasticity of Variance Diffusions. Working paper, Stanford University.

Cox, J., Ingersoll, J. and Ross, S. (1980). An Analysis of Variable Rate Loan Contracts. Journal of Finance 35, 389-403.

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

Davis, R. B. (1987). Hypothesis Testing When a Nuisance Parameter is Present Only under the Alternative. Biometrika, 74, 33-43.

Diebold, F., Lee, J., and Weinbach, G., (1994). Regime Switching with Time-Varying Transition Probabilities, in Hargreaves, C. (ed.), Nonstationary Time Series Analysis and Cointegration, Oxford University Press, Oxford.

Diebold, F., Gunther, T., and Tay, A., (1998). Evaluating Density Forecasts with Applications to Financial Risk Management. International Economic Review, 39. 863–883.

Dothan, U. (1978). On the Term Structure of Interest Rates. Journal of Financial Economics, 6, 59-69.

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

Doukhan, P. (1994). Mixing. Springer-Verlag, New York.

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.

Elerian, O., Chib, S., and Shephard, N. (2001). Likelihood Inference for Discretely Observed Non-linear Diffusions. Econometrica, 69, 959-993.

Engle, R. (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. Econometrica, 50(4), 987-1007.

Forman, J.L. and M. Sørensen (2014) A transformation approach to modelling multi-modal diffusions. *Journal of Statistical Planning and Inference* 146, 56–69.

Gallant, R. and Tauchen, G. (1998). Reprojecting Partially Observed Systems with Application to Interest Rate Diffusions. Journal of American Statistical Association, 93, 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.

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.

Hansen, B. E. (1992). The Likelihood Ratio Test Under Nonstandard Conditions: Testing the Markov Switching Model of GNP. Journal of Applied Econometrics, 7, 61-82.

Hansen, B. E. (1994). Autoregressive Conditional Density Estimation. International Economic Review, 35, 705-730.

Hansen, B. E. (1996). Erratum: The Likelihood Ratio Test Under Nonstandard Conditions: Testing the Markov Switching Model of GNP. Journal of Applied Econometrics, 11(2), 195-198.

Jiang, G., and J. Knight., (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.

Jorion, P. (2006). Value at Risk: The New Benchmark for Managing Financial Risk (3rd ed). McGraw-Hill, New York.

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

Kloeden, P., and Platen, E. (1992). Numerical Solution of Stochastic Differential Equations, Applications of Mathematics. New York: Springer-Verlag.

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

Marsh, T., and Rosenfeld, E. (1983). Stochastic Processes for Interest Rates and Equilibrium Bond Prices. Journal of Finance, 38, 635-646.

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

Patton, A. (2006a). Modelling Asymmetric Exchange Rate Dependence. International Economic Review, 47(2), 527-556.

Patton, A. (2006b). Estimation of Multivariate Models for Time Series of Possibly Different Lengths. Journal of Applied Econometrics, 21, 147-173.

Patton, A. (2009). Copula-Based Models for Financial Time Series. In T.G. Andersen, R.A. Davis, J.P. Kreiss and T. Mikosch, eds. Handbook of Financial Time Series, Springer Verlag, Berlin Heidelberg, 767-785.

Ramsey, J. B. (1969). Tests for Specification Errors in Classical Linear Least Squares Regression Analysis. Journal of the Royal Statistical Society Series B, 31(2), 350–371.

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

Tong, H. (1978). On a Threshold Model, in C.H. Chen (ed.), Pattern Recognition and Signal Processing, Amsterdam: Sijthoff & Noordhoff, 101-41.

Tong, H. (1990). Non-Linear Time Series: A Dynamical Systems Approach, Oxford University Press, Oxford.

Tong, H., and K.S. Lim, (1980). Threshold Autoregressions, Limit Cycles, and Data. Journal of the Royal Statistical Society B, 42, 245-292.

Teräsvirta, T. (1994). Specification, Estimation, and Evaluation of Smooth Transition Autoregressive Models, Journal of the American Statistical Association, 89, 208-218.

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

Table 1: Descriptive statistics of weekly LIBOR and EDR

|                       | LIBOR           | EDR             |
|-----------------------|-----------------|-----------------|
| Sample period         | 1986.01-2007.12 | 1971.01-2007.12 |
| Sample size           | 1148            | 1930            |
| Mean                  | 7.413           | 6.781           |
| Std. Dev.             | 3.214           | 3.550           |
| Skewness              | 1.003           | 1.099           |
| Kurtosis              | 2.872           | 4.809           |
| Jarque-Bera statistic | 193.280         | 651.979         |

Table 2: Maximum likelihood estimation results for LIBOR

| | Model 0 | | Model 1 | | Model 2 | | Model 3 | | Model 4 | | Model 5 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | LIBOR | | | | |
| $\alpha_1$ | -0.096 | $\alpha_1$ | -0.002 | $\alpha_1$ | -0.002 | $\alpha_1$ | -0.002 | $\alpha_1$ | -0.003 | $\alpha_1$ | -0.003 |
| | (0.080) | | (0.006) | | (0.003) | | (0.005) | | (0.001) | | (0.006) |
| $\alpha_2$ | 0.075 | $\alpha_2$ | 0.013 | $\alpha_2$ | 0.003 | $\alpha_2$ | 0.003 | $\alpha_2$ | 0.002 | $\alpha_2$ | 0.001 |
| | (0.067) | | (0.006) | | (0.003) | | (0.002) | | (0.004) | | (0.002) |
| $\sigma^2$ | 0.005 | $\sigma^2$ | 0.005 | $\sigma^2$ | 0.002 | $\sigma^2$ | 0.002 | $\sigma^2$ | 0.001 | $\sigma^2$ | 0.001 |
| | (0.009) | | (0.001) | | (0.001) | | (0.001) | | (0.001) | | (0.001) |
| $\theta$ | 1.330 | $\omega$ | 1.514 | $\theta_L$ | 1.648 | $\theta_L$ | 1.649 | $\theta_L$ | 1.146 | $\theta_L$ | 1.184 |
| | (0.060) | | (0.528) | | (0.070) | | (0.067) | | (0.193) | | (0.190) |
| | | $\beta_1$ | 0.063 | $\theta_H$ | 1.301 | $\theta_H$ | 1.301 | $\theta_H$ | 1.299 | $\theta_H$ | 2.080 |
| | | | (0.044) | | (0.056) | | (0.054) | | (0.053) | | (0.225) |
| | | $\beta_2$ | -0.043 | $\tau_{SET}$ | 12.625 | $\tau_{LST}$ | 12.684 | $p_{LL}$ | 0.877 | $c_L$ | 5.493 |
| | | | (0.004) | | (——) | | (1.460) | | (0.028) | | (11.250) |
| | | $\beta_3$ | 0.003 | | | $\lambda$ | $1.189 \times 10^4$ | $p_{HH}$ | 0.835 | $c_H$ | 30.610 |
| | | | (0.001) | | | | $(1.970 \times 10^6)$ | | (0.019) | | (14.311) |
| | | | | | | | | | | $d_{L_1}$ | -0.726 |
| | | | | | | | | | | | (5.406) |
| | | | | | | | | | | $d_{L_2}$ | 0.055 |
| | | | | | | | | | | | (0.945) |
| | | | | | | | | | | $d_{L_3}$ | -0.007 |
| | | | | | | | | | | | (0.071) |
| | | | | | | | | | | $d_{H_1}$ | -12.868 |
| | | | | | | | | | | | (7.056) |
| | | | | | | | | | | $d_{H_2}$ | 2.111 |
| | | | | | | | | | | | (1.219) |
| | | | | | | | | | | $d_{H_3}$ | -0.152 |
| | | | | | | | | | | | (0.088) |
| lnL | 624.19 | | 671.92 | | 671.20 | | 671.31 | | 1000.80 | | 1019.30 |
| AIC | -1240.39 | | -1327.84 | | -1332.40 | | -1328.62 | | -1987.59 | | -2008.56 |
| BIC | -1220.21 | | -1287.47 | | -1307.17 | | -1293.30 | | -1952.27 | | -1932.87 |
| LR | | | 95.46 | | 94.02 | | 94.24 | | 753.22 | | 790.22 |
| $p$-value | | | 0.0000 | | 0.0000 | | 0.0000 | | 0.0000 | | 0.0000 |

Table 3: Maximum likelihood estimation results for EDR

| | EDR | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | Model 0 | | Model 1 | | Model 2 | | Model 3 | | Model 4 | | Model 5 |
| $\alpha_1$ | -0.079 | $\alpha_1$ | -0.011 | $\alpha_1$ | -0.014 | $\alpha_1$ | -0.011 | $\alpha_1$ | -0.013 | $\alpha_1$ | -0.013 |
| | (0.064) | | (0.020) | | (0.015) | | (0.009) | | (0.005) | | (0.003) |
| $\alpha_2$ | 0.206 | $\alpha_2$ | 0.1273 | $\alpha_2$ | 0.109 | $\alpha_2$ | 0.108 | $\alpha_2$ | 0.047 | $\alpha_2$ | 0.044 |
| | (0.170) | | (0.072) | | (0.107) | | (0.086) | | (0.019) | | (0.020) |
| $\sigma^2$ | 0.027 | $\sigma^2$ | 0.035 | $\sigma^2$ | 0.034 | $\sigma^2$ | 0.034 | $\sigma^2$ | 0.012 | $\sigma^2$ | 0.012 |
| | (0.002) | | (0.003) | | (0.004) | | (0.004) | | (0.001) | | (0.002) |
| $\theta$ | 1.219 | $\omega$ | 1.199 | $\theta_L$ | 1.123 | $\theta_L$ | 1.123 | $\theta_L$ | 1.074 | $\theta_L$ | 1.078 |
| | (0.026) | | (0.705) | | (0.030) | | (0.030) | | (0.028) | | (0.038) |
| | | $\beta_1$ | 0.519 | $\theta_H$ | 1.268 | $\theta_H$ | 1.268 | $\theta_H$ | 1.659 | $\theta_H$ | 1.688 |
| | | | (0.126) | | (0.029) | | (0.029) | | (0.053) | | (0.066) |
| | | $\beta_2$ | -0.074 | $\tau_{SET}$ | 11.610 | $\tau_{LST}$ | 11.619 | $p_{LL}$ | 0.967 | $c_L$ | 1.095 |
| | | | (0.020) | | (——) | | (0.121) | | (0.007) | | (2.232) |
| | | $\beta_3$ | 0.003 | | | $\lambda$ | 999.997 | $p_{HH}$ | 0.933 | $c_H$ | 2.572 |
| | | | (0.002) | | | | ($5.909\times10^3$) | | (0.018) | | (2.047) |
| | | | | | | | | | | $d_{L_1}$ | 1.023 |
| | | | | | | | | | | | (1.342) |
| | | | | | | | | | | $d_{L_2}$ | -0.093 |
| | | | | | | | | | | | (0.258) |
| | | | | | | | | | | $d_{L_3}$ | -0.001 |
| | | | | | | | | | | | (0.019) |
| | | | | | | | | | | $d_{H_1}$ | 0.413 |
| | | | | | | | | | | | (1.088) |
| | | | | | | | | | | $d_{H_2}$ | -0.097 |
| | | | | | | | | | | | (0.195) |
| | | | | | | | | | | $d_{H_3}$ | 0.005 |
| | | | | | | | | | | | (0.014) |
| lnL | 425.09 | | 443.91 | | 445.07 | | 445.07 | | 862.88 | | 873.41 |
| AIC | -842.18 | | -871.82 | | -880.14 | | -876.14 | | -1711.76 | | -1716.82 |
| BIC | -819.92 | | -827.30 | | -852.31 | | -837.18 | | -1672.80 | | -1633.34 |
| LR | | | 37.64 | | 39.96 | | 39.96 | | 875.58 | | 896.64 |
| $p$-value | | | 0.0000 | | 0.0000 | | 0.0000 | | 0.0000 | | 0.0000 |

Figure 1: Original and differenced series of weekly LIBOR and EDR
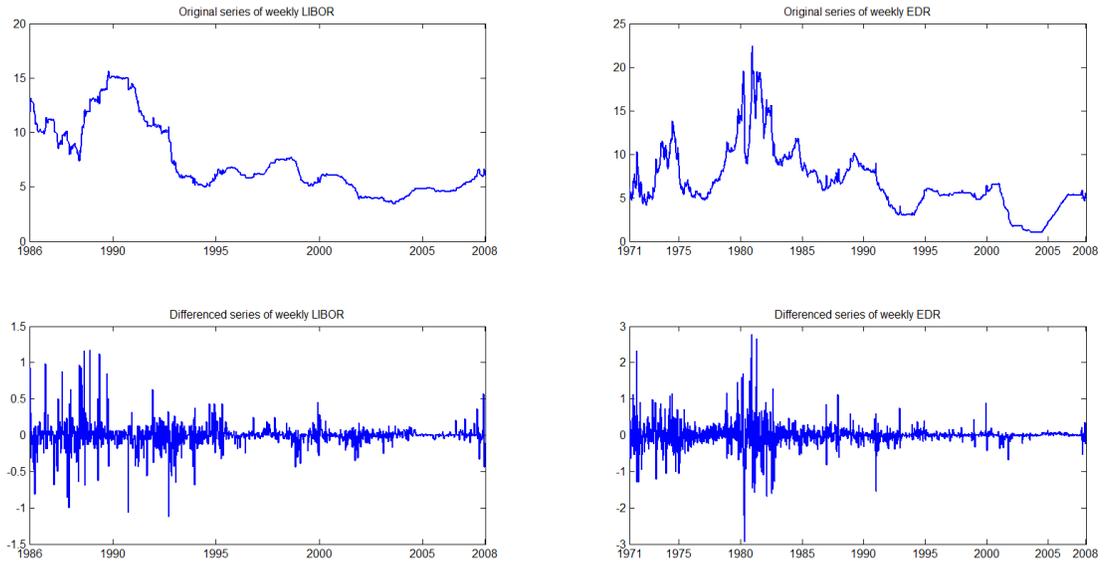
Figure 2: Estimated time-varying transformation parameter from Model 1
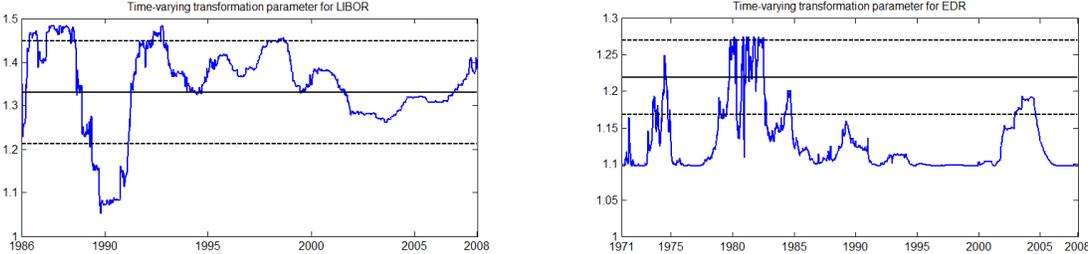
Figure 3: Smoothed probabilities from Model 4
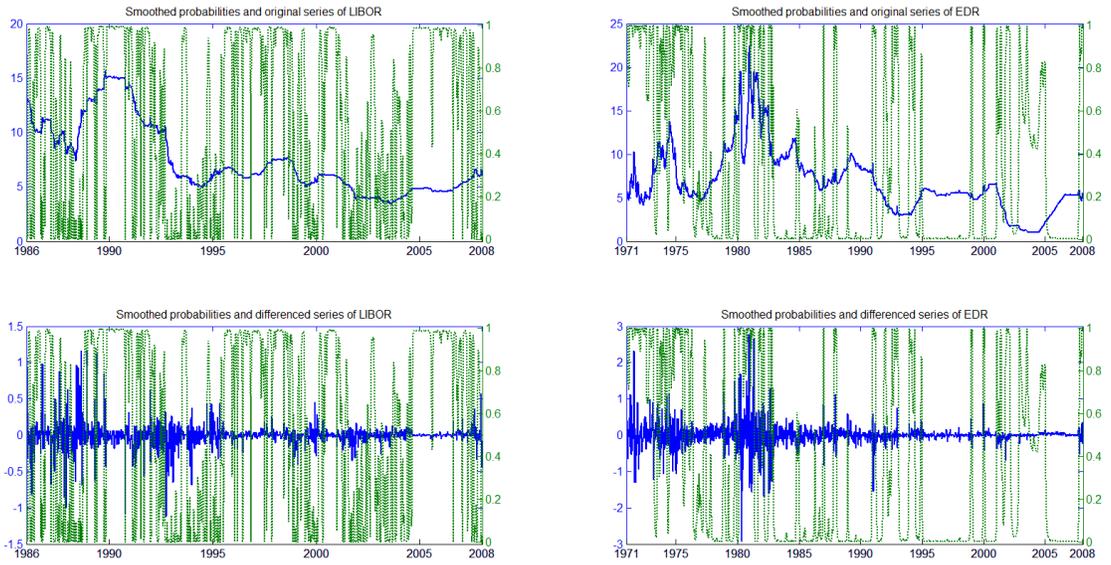
Figure 4: Estimated endogenous time-varying transition probabilities from Model 5