2017

Calibration of vasicek model in a hidden markov context: the case of Kenya

John Kigen Chelimo
Strathmore Institute of Mathematical Sciences (SIMs)
Strathmore University

Calibration of Vasicek Model in a Hidden Markov Context: The Case of Kenya

Chelimo, John Kigen

Submitted in partial fulfillment of the requirements for the Master of Science in Mathematical Finance at Strathmore University

Strathmore Institute of Mathematical Sciences

June 2017

Abstract

This dissertation calibrates the Vasicek term-structure model to the evolution of interest rate dynamics in Kenya. This is done for both a single-state and a multi-state model using state estimated under a Hidden Markov Model (HMM). The findings of this paper provide a starting point for the management of the risk posed by interest rate-dependent instruments.

The Vasicek model is calibrated using monthly observations of the 91-day Treasury bill rate from September 1994 to July 2014 as a proxy for the short rate. Key results show an increase in the mean reversion parameter with an increase in the number of states, suggesting higher stability of states. The volatility is observed to move independently of the level of the interest rate, supporting the idea that risk is not necessarily a function of the level of the interest rate but rather related to the inherent variability of rates in a particular state. Findings from this parameter estimation provide support for interest rate models that incorporate regime switches.

Keywords: Vasicek model, hidden Markov model(HMM), regime-switches, calibration

Declaration

I declare that this work has not been previously submitted and approved for the award of a degree by this or any other University. To the best of my knowledge and belief, the dissertation contains no material previously published or written by another person except where due reference is made in the dissertation itself.

Chelimo, John Kigen

Signed: ....

June 2017

Approval

This dissertation of Chelimo, John Kigen was reviewed and approved by the following:

Dr. Samuel Chege Maina,
Lecturer, Strathmore Institute of Mathematical Sciences,
Strathmore University

Ferdinand Othieno,
Dean, Strathmore Institute of Mathematical Sciences,
Strathmore University

Professor Ruth Kiraka,
Dean, School of Graduate Studies,
Strathmore University

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Overview

One of the most important concepts in finance is the time value of money; the idea that cash today is worth more than a similar amount of cash in future. The measure that allows assignment of value across time is the interest rate and its application is critical in every aspect involving intertemporal transfer of funds.

Interest rate is the compensation you would require to forego consumption today in return for a higher amount in future. As this concept is so well accepted, and its application so frequent, it may seem that no formal treatment is required. However, to determine the dynamics of interest rate-dependent financial variables, or to manage the risk exposure presented by such assets, a completely different set of challenges is faced that cannot be left to anecdotal evaluation. Indeed, Cuchiero (2006) notes that although the concept of interest rates is so pervasive, managing interest rate risk presents a distinct level of complexity that is best dealt with in formal mathematical modeling.

Given the importance of understanding interest rate dynamics and the extensive areas of application, there is a significant body of knowledge developed covering the theory and practice of interest rate modeling. However, this extensive literature has not extended to empirical work in the Kenyan market. This research focuses on the modeling of interest rate dynamics in the context of regime switches. The goal is to provide an empirical evaluation of the behaviour of short-rate model parameter estimates in the context of regime switches.

1.2 Background to the study

1.2.1 Overview of interest rate modeling

Models of interest rate dynamics are integral to the pricing of fixed income and derivative instruments. Mathematical interest rate models have stemmed from the need to adequately model and manage the risk presented by future movements in interest rates.

As a result, various interest rate models have been developed in an attempt to answer the questions of; 1) which quantities and dynamics should be modeled? 2) How should their randomness be modeled? and 3) what the valuation consequences of the different approaches are (Brigo, 2007). Interest rate modeling theory has followed four key veins in an attempt to answer these questions:

1. Modeling the short rate
2. Modeling the instantaneous forward rates
3. Modeling of forward market-rates
4. Market rate modeling with volatility smile extensions

While the models have different levels of complexity, all formulations have different applications including advanced pricing, rating practice as well as risk management. (Brigo, 2007).

This research focuses on short rate models, particularly the Vasicek model which captures the mean-reversion element of interest rates while retaining analytical tractability. Other short-rate models have been developed which include the Cox-Ingersoll-Ross model, the Dothan model, the Black-Derman-Toy model, the Ho-Lee model, and the Hull-White (Extended Vasicek) model. These models remedy some of the weaknesses of the Vasicek model, including the possibility of negative rates in the Vasicek model, but lose the computational tractability offered by use of a Gaussian distribution. On the other hand, models of the instantaneous forward rate, and particularly the model by Heath et al. (1992) capture the full dynamics of the yield curve, but are non-Markovian. This makes the model computationally intractable, and reduces its attractiveness for this study. Finally, market rate models present a good class of models that allow direct modeling of observable bond prices, as opposed to unobserved variables used in the short-rate and forward rate models. In the Kenyan context, quality data on bond prices is a challenge, which limits the use of bond-price based models. It is also observed that the additional complexity of the LIBOR market models, while increasing the pricing accuracy, is less relevant for the Kenyan context where the spectrum of available interest rate derivative instruments remains narrow and fairly unsophisticated.

Short-rate term structure modeling

Under short-rate term-structure modeling, the interest rate, $r(t)$ is modeled as a stochastic differential equation with a drift and a diffusion component. Each short-rate model specifies a different evolution for the dynamics of $r(t)$. Risk management is most suitably addressed by tractable short-rate models due to computational ease, which makes them attractive for a large number of firms. In contrast, pricing models require higher precision in the distribution which is not provided by short-rate models. (Brigo, 2007). Short-rate modeling is discussed in greater detail in subsequent sections.

Forward rate modeling

These models incorporate more parameters and were developed to accommodate more flexible option structures as well as give less-correlated rates at future times. (Brigo, 2007)

Following Brigo (2007), the market-based forward LIBOR at time t between T and S is defined as,

$$F(t;T,S) = \frac{\frac{P(t,T)}{P(t,S)} - 1}{S - T}. \tag{1.1}$$

In the limit as S tends to T, the instantaneous forward rate is given as:

$$f(t,T) = \lim_{T \to S} F(t;T,S) \tag{1.2}$$

$$= -\frac{\partial lnP(t,T)}{\partial T}, f(t,T) = r(t) \tag{1.3}$$

Given the flexible dynamics of r, there are restrictions to the dynamics allowed for f. This is a fundamental theoretical result due to Heath et al. (1992), more commonly referred to as HJM. If no arbitrage is to hold, and by setting $f(0,T) = f^{Market}(0,T)$, under the risk-neutral measure:

$$df(t,T) = \sigma(t,T)\left(\int_t^T \sigma(t,s)ds\right)dt + \sigma(t,T)dW(t), \tag{1.4}$$

is obtained.

This relation links the local mean to the local standard deviation by the no-arbitrage property of interest rate dynamics. This is different from models based on dr_t, where the whole risk neutral dynamics was free.

The Heath et al. (1992) model serves as a useful framework by which various no-arbitrage interest-rate models can be unified. In practice however, any useful models coming out of the HJM are either the short-rate models or Market models.

Modeling of Forward Market rates

Modeling for LIBOR market models (LMMs) involves outlining families of forward rates, F_i spanning i associated with the relevant forward rate agreements rather than modeling either r or f.

The forward LIBOR rate between T_{i-1} and T_i, at time t is defined as:

$$F_i(t) = \frac{\frac{P(t,T_{i-1})}{P(t,T_i)} - 1}{(T_i - T_{i-1})} \tag{1.5}$$

The LMM is compatible with Black's market formula and provides a rigorous arbitrage free justification for the formula. The quantities in the LIBOR market models are structured because they are forward rates coming from expectations of objects involving r. In particular, the LIBOR market model was developed starting from instantaneous forward-rate-dynamics in Brace, Gatarek and Musiela (1997), although it is possible to obtain it also through the change-of-numeraire approach. (Brigo and Mercurio, 2006)

Volatility smile extensions of Forward Market models

This recent extension of interest rate modeling involves incorporation of volatility smile effects. The volatility smile effect applies extensively to option contracts and is expressed in terms of market rates volatilities. Therefore, models incorporating volatility smiles are market models rather than short-rate or forward-rate models.

Brigo (2007) outlines two main approaches to accommodate the volatility smile effects:

- *Local volatility models*: Here σ^2 is specified as a function of the underlying. Local volatility models include the Constant Elasticity of Variance (CEV) model and the Displaced diffusion model. As there is no new randomness added to the system as time moves on, the models imply a volatility smile that flattens in time.

- *Stochastic volatility models*: Under this family of models, σ^2 is specified as a new stochastic process, adding new randomness to the volatility. Consequently, volatility becomes a variable with a new random life of its own that could be correlated with the underlying. Stochastic volatility models include the Heston Stochastic Volatility model (1993) or the more simplistic and popular Stochastic Alpha Beta Rho (SABR) model (2002)

1.2.2 Interest rate modeling in Kenya

Interest rate modeling in Kenya so far has focused on more descriptive approaches. Few models have focused on evaluating the interest rate models that incorporate the empirical dynamics of the interest rate in Kenya.

In analysing interest rate dynamics in Kenya, Caporale and Gil-Alana (2010) examine the stochastic properties of the interest rate spread to provide useful information about the effects of shocks and appropriate policy responses. They test the stationarity of the interest rate series. They reject the hypothesis of mean reversion in the interest rate process. They estimate the orders of integration to be equal to or higher than 1 in all cases. For interest rate spreads, mean reversion is only found in the case of the deposits-Treasury bill rate spread under the assumption of autocorrelated errors.

To understand interest rate dynamics, Ngugi (2001) focuses on the factors influencing the interest rate spread. Her work focused on the dynamics of the interest rate spread over time and she demonstrates that the spread changes significantly, in line with policy as well as economic conditions. This provides support for an evaluation of interest rate dynamics that incorporates regime switches.

Olweny (2011) focuses on the link between short-term volatility of the interest rate and the level of interest rates in Kenya using the Treasury bill rates from August 1991 to December 2007. His findings indicate that the volatility is positively correlated with the level of the short-term interest rate. He also finds that the GARCH model is better suited for modeling volatility of short rates in Kenya, compared to ARCH models.

1.3 Problem statement

Modeling and estimating the dynamics of interest rates is critical in the pricing of bonds, options and other derivatives. While there is a large volume of interest rate modeling theory, very little has been done to evaluate the performance of such models in Kenya, particularly under regime switching.

This research aims to contribute to both the theory and practice of interest rate modeling in Kenya. This evaluation will also involve incorporation of regime-switches to the Vasicek model in line with changing underlying economic variables.

The results of this analysis form a good foundation for the pricing and risk management for interest rate dependent securities and derivatives. In this way, this research provides both theoretical as well as practical insights to the interest rate risk management in Kenya.

1.4 Research objectives

This dissertation studies questions related to modeling of the short-rate in Kenya. The research seeks to answer the primary question of what the parameters of a good model of the short-term interest rate should be in the context of regime switches.

1.4.1 Specific objectives

The main goals of this dissertation can be summarised as:

- *Parameter estimation for the Vasicek model.*: What are the parameters of the Vasicek model when applied to Treasury bill rates in the Kenya?
- *Evaluation of impact of regime switches on parameter estimates.*: What is the behaviour of the estimated parameters when regime switching ia introduced in modeling the underlying interest rate process?

Chapter 2

Literature review

2.1 Overview of interest rate theory

This section provides an overview of key definitions and concepts in interest rate modeling. Following the approach used in Cuchiero (2006), key concepts on the short-term interest rate, zero-coupon bond and no-arbitrage pricing are expounded upon. Additional derivations and relationships are available in Brigo and Mercurio (2006), Bjork (2009) and Musiela and Rutkowski (2005).

2.1.1 Definition: Short-term interest rate

A risk-free security, compounded continuously at a risk-free considered to be the instantaneous / short-term interest rate is defined as follows:

Let r_t denote the rate for risk-free borrowing or lending at time t over the infinitesimal time interval $[t, t+dt]$. The rate r_t is assumed to be an adapted process on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}(\mathrm{t})_{0<\mathrm{t}<T^*})$ for some $T^* > 0$[2], with almost all sample paths integrable on $[0, T^*]$. $B(t) = B(t, \omega)$ is defined to be the value of the bank account at $t \geqslant 0$ that evolves for almost all $\omega \in \Omega$ according to the differential equation:

$$dB(t) = r(t)B(t)dt, \text{with } B(0)\text{=}1 \tag{2.1}$$

Consequently:

$$B(t) = exp\left(\int_0^t r(s)ds\right) \text{for all } t \in [0, T^*] \tag{2.2}$$

Therefore, B_t allows the relation of two amounts of currency available at different points in time.

Stochastic discount factor

The stochastic discount factor $D(t,T)$ is the value at time t of one unit of cash payable at time $T > t$ and is given by:

$$D(t,T) = \frac{B(t)}{B(T)} = exp\left(-\int_t^T r(s)ds\right) \tag{2.3}$$

2.1.2 Definition: Zero coupon bond

A zero-coupon bond of maturity T is a financial security paying one unit of cash at a pre-specified date T in the future without intermediate payments. The price at time $t < T$ is denoted by $P(t,T)$. $P(T,T) = 1$.

Note that there is a close relationship between the zero-coupon bond price $P(t,T)$ and the stochastic discount factor $D(t,T)$. Actually, $P(T,T)$ corresponds to the expectation of $D(t,T)$ under the risk neutral probability measure. (Cuchiero, 2006)

2.2 The no-arbitrage term-structure equation

The absence of arbitrage opportunities between all bonds with different maturities and the bank account is the fundamental economic assumption informing interest rate theory.

A family $P(t,T)$, $t < T < T^*$ of adapted processes is called an arbitrage-free family of bond prices relative to r if the following conditions hold:

1. $P(T,T) = 1$ for all $t \in [0,T^*]$, and
2. There exists a probability measure $\mathbb{P}^*$ on $(\Omega, \mathcal{F}_{T^*})$ equivalent to $\mathbb{P}$, such that for all $t \in [0,T]$, the discounted bond price:

$$\tilde{P}(t,T) = D(0,T)P(t,T) = \frac{B(0)}{B(t)}P(t,T) = \frac{P(t,T)}{B(t)} \tag{2.4}$$

   is a martingale under $\mathbb{P}^*$.

A probability measure $\mathbb{P}^*$ that satisfies the conditions of the definition above is called a martingale measure for the family $P(t,T)$. This definition is based on the general result that arbitrage opportunities are absent given the existence of an equivalent martingale measure in a standard market model. Given that $P(t,T)$ follows a martingale under $\mathbb{P}^*$:

$$\tilde{P}(t,T) = \mathbb{E}_{\mathbb{P}^*}\left(\tilde{P}(T,T)|\mathcal{F}_t\right) \text{ for } t \leq T. \tag{2.5}$$

Therefore,

$$D(0,T)P(t,T) = \mathbb{E}_{\mathbb{P}^*}\left(D(0,T)P(T,T)|\mathcal{F}_t\right) = \mathbb{E}_{\mathbb{P}^*}\left(D(0,T)|\mathcal{F}_t\right), \tag{2.6}$$

Which leads to the following expression for the bond: price:

$$\begin{aligned} P(t,T) &= D(0,T)^{-1}\mathbb{E}_{\mathbb{P}^*}\left(D(0,T)|\mathcal{F}_t\right) & (2.7)\\ &= exp\left(-\int_t^T r(s)ds\right)\mathbb{E}_{\mathbb{P}^*}\left(exp\left(-\int_t^T r(s)ds\right)|\mathcal{F}_t\right) & (2.8)\\ &= \mathbb{E}_{\mathbb{P}^*}\left(exp\left(-\int_t^T r(s)ds\right)|\mathcal{F}_t\right) = \mathbb{E}_{\mathbb{P}^*}\left(D(t,T)|\mathcal{F}_t\right) & (2.9) \end{aligned}$$

Thus $P(t,T)$ corresponds to the expectation of the stochastic discount factor $D(t,T)$ under $\mathbb{P}^*$. The no-arbitrage bond prices are directly obtained. These are a special case of the general no-arbitrage prices associated with an attainable contingent claim H given by:

$$\pi_t = \mathbb{E}_{\mathbb{P}^*}\left(D(t,T)H|\mathcal{F}_t\right) \tag{2.10}$$

Where it is assumed that the price process $P(t,T)$ follows a strictly positive and adapted process on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{0\leq t\leq T^*})$, where the filtration $\mathcal{F}_t$ is again the $\mathbb{P}$-completed version of the filtration generated by the underlying Brownian motion. (Cuchiero, 2006)

2.3 Key literature

2.3.1 Term structure models

Term structure modeling in continuous time lends itself to various approaches. Most frequently, the short term interest rate is assumed to follow a diffusion process. Bond prices are determined as solutions to a partial differential equation which places restrictions on the relationship of risk premia of bonds with different maturities. The challenge with this approach is that it is particularly difficult and cumbersome to fit the observed term structure of interest rates within the simple diffusion model. (Svoboda, 2002).

The Vasicek (1977) model was the first to make a significant impact on interest rate modeling. The model makes assumptions about the stochastic evolution of interest rates by exogenous specification of the short-term interest rate process. The Vasicek and most other models are partial equilibrium models. Beliefs about future realizations of the short-term interest rate are taken as inputs and used to make assumptions about investor preferences (as specified by the market prices of risk).

A later approach used by Cox et al. (1985) starts with the specification of an equilibrium economy which forms the foundation for specification of the model. The model makes assumptions about the stochastic evolution of exogenous state variables and investor preferences. The equilibrium economy provides for endogenous derivation of the form of the short rate and hence the prices of contingent claims. Production opportunities, investor tastes and beliefs about future states of the world provide an exogenous specification of the economy from which bond prices are derived, to make the CIR a complete equilibrium model.

The relationship between interest rates and bonds with different maturity times is given by the term structure of interest rates. For this research, we calibrate the Vasicek model which directly models the dynamics of the instantaneous short rate $r(t)$. The annualized interest rate for an infinitesimally short time period is taken to be the short rate, but in practice the three-month rate is considered a better approximation of the short-rate. (Van Elen, 2010). This research uses the 91-day Treasury bill rate as a proxy for the short-rate.

Following Van Elen (2010), the short rate is defined as:

$$r(t) = R(t, 0) = \lim_{T \to 0} R(t, T), \tag{2.11}$$

where t denotes a moment in time, T, the time to maturity and $R(t, T)$ the corresponding interest rate. Let $P(t, T)$ denote the value of a zero-coupon bond at time t that pays 1 at maturity time T, and $R(t, T)$ the corresponding interest rate. In continuous time, this is found to be:

$$P(t,T) = exp\left(-R(t,T)(T-t)\right) \tag{2.12}$$

Rewriting the equation above yields a way to describe the interest rate as a function of the value of a bond:

$$R(t,T) = -\frac{1}{T-t} log\left(P(t,T)\right). \tag{2.13}$$

Since $P(t,T)$ is a simple discount factor, it is clear that $R(t,T)$ may not be negative to make sure the discount factor as described in $P(t,T)$ lies between 0 and 1.

2.3.2 Vasicek model

The Vasicek model assumes that the instantaneous spot rate under the real-world measure evolves as an Ornstein-Uhlenbeck(O-U) process with constant coefficients. This is equivalent to assuming that r follows an O-U process with constant coefficients under the risk-neutral measure for a suitable choice of the market price of risk, t. Vasicek (1977) defines the short rate process as:

$$dr(t) = \kappa[\theta - r(t)]dt + \sigma dW(t), r(0) = r_0 \tag{2.14}$$

where r_0,κ, and σ are positive constants.
Integrating, for each $s < t$:

$$r(t) = r(s)e^{-\kappa(t-s)} + \theta(1 - e^{-\kappa(t-s)} + \sigma \int_s^t e^{-\kappa(t-u)} dW(u), \tag{2.15}$$

so that $r(t)$ conditional on $\mathcal{F}_s$ has a normal distribution with mean and variance given respectively by:

$$\mathbb{E}[(r(t)|\mathcal{F}_s] = r(s)e^{-\kappa(t-s)} + \theta(1 - e^{-\kappa(t-s)}) \tag{2.16}$$

$$Var[r(t)|\mathcal{F}_s] = \frac{\sigma^2}{2\kappa}[1 - e^{2\kappa(t-s)}], \tag{2.17}$$

A major downside of the Vasicek model is that for each time t, the rate $r(t)$ can be negative with positive probability. Despite this, the analytical tractability that is implied by a Gaussian density is hardly achieved when assuming other distributions for the process r.

Given the dynamics of the Vasicek model, the maximum-likelihood estimator for the Vasicek model is now considered. The dynamics are expressed as:

$$dr(t) = [b - ar(t)]dt + \sigma dW^0(t) \tag{2.18}$$

with b and a as suitable constants. By integration between any instants s and t, $r(t)$ conditional on $\mathcal{F}_s$ is normally distributed with mean $r(s)e^{-a(t-s)} + \frac{b}{a}\left(1 - e^{-a(t-s)}\right)$, and variance $\frac{\sigma^2}{2a}\left[1 - e^{-2a(t-s)}\right]$.

The following functions of the parameters are estimated:

$$r(t) = r(s)e^{-a(t-s)} + \frac{b}{a}\left(1 - e^{-a(t-s)}\right) + \sigma \int_s^t e{-}a(t-u)dW^0(u) \tag{2.19}$$

Calibration of parameters can be done using the Ordinary Least Squares (OLS) method or by Maximum Likelihood Estimation(MLE). The MLE estimators for α, β, σ are derived as:

$$\hat{\alpha} = \frac{\sum_{i=1}^{n} r_i, r_{i-1} - \sum_{i=1}^{n} r_i \sum_{i=1}^{n} r_{i-1}}{n \sum_{i-1}^{n} r_{i-1}^2 \left(\sum_{i-1}^{n} r_{i-1}\right)^2} \tag{2.20}$$

$$\hat{\beta} = \frac{\sum_{i-1}^{n} r_i - \hat{\alpha} r_{i-1}}{n(1-\hat{\alpha})} \tag{2.21}$$

$$\hat{V^2} = \frac{1}{n}\sum_{i=1}^{n}\left[r_i - \hat{\alpha} r_{i-1} - \beta(1-\hat{\alpha})\right]^2 \tag{2.22}$$

The estimated quantities give full information on the transition probability for the process r under Q_0. This allows for simulations at one-day spaced future time instants.

2.4 Modeling with regime switches

Most interest rate models, including those specified by Vasicek (1977) and Cox et al. (1985), assume constant parameters over the relevant sample period and are thus single-regime models. It is however likely that, over time, the economic and political conditions that generate interest rates may change. In such cases, the parameters of a model of the interest rate, or even the model structure itself may change. Dahlquist and Gray (1998).

In particular, Wu and Zeng (1991) find that regimes underlying interest rate term structures are related to business cycles. The risk of regime shifts is therefore likely a systematic risk.

Given that the stochastic behaviour of interest rates varies over time, regime switching (RS) models constitute an attractive class of models to incorporate the stochastic behavior of interest rates within a stationary model.(Ang and Bekaert, 2001)

Short-term interest rates have two important empirical attributes.

- First is the mean reversion property of the process, modeled by letting the next period's change in the short rate depend linearly on the current level of the short rate.
- Second, is that the unconditional distribution of changes in the short rate is leptokurtic.

These two effects are captured by considering a class of models that are based on various continuous time/diffusion models, and extending this class of models to allow for regime shifts.(Dahlquist and Gray, 1998).

This research uses the generalized regime-switching (GRS) model. The GRS model nests the GARCH (1,1) model, a discretized diffusion model motivated by Cox et al. (1985) model, and the Markov regime-switching model. (Gray, 1996)
The GRS model is generalized in the sense that:

- Each regime reverts at a different rate to a different long-run mean;
- Conditional variance in each regime takes a very general form incorporating level effects and GARCH effects consistent with a square root process; and
- Switching probabilities are time-varying, depending on the level of the short rate.

The basic premise of the GRS model is that, depending on the latent regime indicator, the parameters of the conditional mean and conditional variance process are allowed to take different values. (Gray, 1996).

Chapter 3

Methodology

3.1 Estimation of parameters of term-structure models

Two alternatives for estimating the term-structure model parameters are considered. The first option is to carry out a cross-sectional estimation where parameter estimation is made at a fixed moment in time while considering the different maturity times. The second alternative is a time-series estimation, where maturity time is fixed and parameters are estimated while considering the evolution of the interest rate over the different time periods in the dataset. (Van Elen, 2010).

If one-factor term-structure models were true, then the difference between time series and cross-sectional estimates should be small. Van Elen (2010) shows that while models based purely on the short-rate may not be very realistic in all scenarios, they might still do a good job in term-structure fitting.

The analysis in this dissertation uses monthly 90-day treasury bills as a proxy for the instantaneous short rate.

3.2 Model calibration and parameter estimation

3.2.1 Calibration of the Vasicek model

To estimate the parameters, a discrete time version of the model is needed. Following Van Elen (2010), the discrete version of the Vasicek model presented in Brigo (2007) is considered.

$$r(t_i) = c + br(t_{i-1}) + \sigma Z, \tag{3.1}$$

where:

$$c = \mu \left(1 - e^{-\theta \Delta t}\right), \tag{3.2}$$

$$b = e^{-\theta \Delta t}, \tag{3.3}$$

$$\delta = \sigma \sqrt{\frac{(1 - e^{-2\theta \Delta t})}{2\theta}} \tag{3.4}$$

Here Z follows a standard normal distribution, the parameters θ, μ, and σ denote the parameters of the continuous Vasicek model and $\Delta t = t(i) - t(i-1)$.

Parameter calibration is carried out by an ordinary least squares (OLS) regression, providing maximum likelihood estimators for the parameters c,b, and δ. According to Brigo et al. (2008) the following expressions for θ, μ, and σ, hold:

$$\theta = \frac{-log(b)}{\Delta t}, \tag{3.5}$$

$$\mu = \frac{c}{1-b}, \tag{3.6}$$

$$\sigma = \frac{\delta}{\sqrt{\frac{(b^2-1)\Delta t}{2log(b)}}}. \tag{3.7}$$

The values for θ, μ, and σ are derived directly by estimators via maximum likelihood.

$$\hat{b} = \frac{n\sum_{i=1}^{n} r_i r_{i-1} - \sum_{i=1}^{n} r_{i-1}}{n\sum_{i-1}^{n} r_{i-1}^2 - \left(\sum_{i=1}^{n} r_{i-1}\right)^2}, \tag{3.8}$$

$$\hat{\mu} = \frac{\sum_{i-1}^{n}\left[r_i - \hat{b}r_{i-1}\right]}{n\left(1-\hat{b}\right)}, \tag{3.9}$$

$$\hat{\delta}^2 = \frac{1}{n}\sum_{i=1}^{n}\left[r_i - \hat{b}r_{i-1} - \hat{\mu}\left(1-\hat{b}\right)\right]^2. \tag{3.10}$$

With these estimators and equations, estimators for θ and σ are readily found.

3.2.2 Hidden Markov model (HMM) filtering

Analysis of the sample data indicates that the evolution of the Treasury bill rate undergoes several distinct regimes characterized by states with high and low means as well as high and low standard deviations. The regime-switching model is proposed to capture such behaviour.

To determine the regimes for the interest rate, a Hidden Markov Model (HMM) is used. Under the HMM, a prior state segregation of states is not required, although, the number of states needs to be determined. The parameter estimates under HMM are calculated through the expectations maximization (EM) algorithm which requires initial values for the implementation. For this purpose, the initial values are found by employing a least-square method on the first few data points.(Erlwein, 2008)

Hidden Markov Models are typically used to predict the hidden regimes of observation data. Therefore, this model has found extensive applications in areas such as speech recognition systems, molecular biology and financial markets. In particular, it has been used in financial markets to determine underlying economic regimes.(Erlwein, 2008).

For the case of interest rate modeling, there are a number of studies that have used Hidden Markov Models as the basis for determining regime switches. Elliott and Wilson (2007) use a hidden Markov Model in the context of the Canadian term structure of interest rates. They find that a three-state

Markov model provides a good framework for explaining data over a basic model that excludes the Markov chain. Maheu and Yang (2015) develop an infinite hidden Markov model for the short-term interest rates. The model provides significant improvements in density forecasts as well as point forecasts. They find evidence of recurring regimes as well as structural breaks in the empirical application. Munclinger (2011) explores the Brazilian term structure in a hidden Markov framework. He applies a hidden Markov model of the term structure to modeling the Brazilian swap rate curve. From this, two regimes are identified, and are characterized as persistent, and are explained by the level and the slope of the term structure.

The HMM assumes that a Markov chain is embedded in the stochastic interest rate process. The Markov chain itself is not observable, i.e., it is ”hidden” in the observations. The aim in HMM filtering is to estimate the underlying Markov chain and this is done by filtering the sequence x_k out of the observations.(Erlwein, 2008). The underlying Markov chain x_k is assumed to be homogenous with finite state space in discrete time.(Erlwein, 2008).

Under the real world measure $\mathbb{P}$, the Markov chain follows the dynamics:

$$x_{k+1} = \pi x_k + v_{k+1} \tag{3.11}$$

where π is the transition probability matrix and v_{k+1} is a martingale increment. The observation process is denoted by y_k and can follow various types of dynamics.(Erlwein, 2008)

For the Vasicek model, the observation process $y_k : k\epsilon\mathbb{N}$ is assumed to have the form:

$$y_{k+1} = \alpha(r_k)y_k + \kappa(r_k) + \xi(r_k)\omega_{k+1} \tag{3.12}$$

The filtrations generated by the process are defined as $\mathcal{F}^v = \sigma(y_1, y_2, ...), \mathcal{F}^x = \sigma(x_1, x_2, ...)$ and $\mathcal{F} = \mathcal{F}^v \bigvee \mathcal{F}^x$. The derivation of optimal parameter estimates is carried out under a reference probability measure under which the observation process and the Markov Chain and independently and identically distributed processes (IID)

Selecting the number of states for the HMM-based interest rate model can be done with a penalized likelihood criteria. The optimal number of regimes for the proposed HMM-based interest rate model is determined by applying the Akaike Information Criterion (AIC). The derivation of this model selection criteria is based on the Kullback-Leibler information and utilizes the log-likelihood function of the model together with the number of model parameters as the penalty term.(Erlwein, 2008)

Chapter 4

Data analysis and results

4.1 Data description

This research uses the 91-day Treasury bill rate as a proxy for the instantaneous short rate to estimate the parameters of a single-regime and multiple-regime Vasicek model. The sample is a dataset of 239 monthly observations covering the period from September 1994 to July 2014. The data is provided by the Central Bank of Kenya.

4.1.1 Motivation for different regimes

From the evolution of the 91-day Treasury-bill rate, significant changes are observed in average level as well as the volatility of the process. The time-series plot for the 91-day Treasury bill rate is shown in figure 4.1 below. Preliminary analysis of the interest rate dynamic shows that the interest rate model parameters are likely to have changed with different states of the economy.

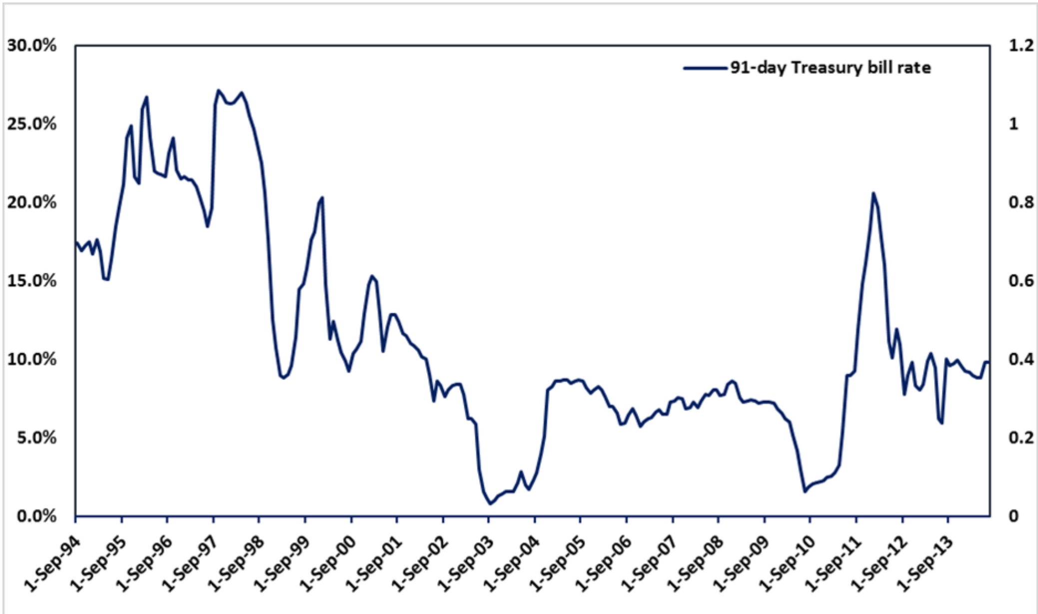

Figure 4.1: Evolution of the 91-day Treasury bill rate from September 1994 to July 2014

Over the entire period, two distinct regimes are observed in the data. The first regime has a high average interest rate which starts in September 1994 through to the end of year 2000. Subsequently, the 91-day rate takes a lower average level. The two regimes point to significant changes in the underlying process driving interest rates. This is partially explained by the start of a new political regime at the end of 2002 and particularly by changes in underlying economic variables.

4.2 Determination of regimes

The hidden Markov model is used to determine the different states over the sample period. Implementation of a 2-state, 3-state and 4-state HMM filter is done and the results observed are shown below.

4.2.1 2-regime model

The HMM fitting procedure for a 2-state model converges at 21 iterations with a LogLikelihood of 457.7127 and 7 degrees of freedom. Information criteria are obtained as -901.4255 for the AIC and -877.0902 for the BIC. The 2-state model starts with the assumption that the interest rate process starts in state 2. Based on the HMM filtering, the transition probability matrix in figure 4.2 below is obtained:

	To State 1	To State 2
From State 1	0.988	0.012
From State 2	0.048	0.952

Figure 4.2: Transition probability matrix for the 2-state HMM model

State 1 has the highest same-state transition probability with a lower value observed for state 2. This suggests higher stability for state 1 than state 2.

By assuming the interest rate process is Gaussian, the descriptive statistics observed are shown in figure 4.3 below:

Response parameters	Re1. (Intercept)	Re1. Sd
State 1	0.076	0.031
State 2	0.207	0.038

Figure 4.3: Descriptive statistics for the 2-state HMM model

It is observed that state 1 has a low long-term mean, and a low standard deviation, while state 2 has a higher mean and standard deviation.

The graphical representation of state changes for the 2-state HMM model are as shown in figure 4.4 below:

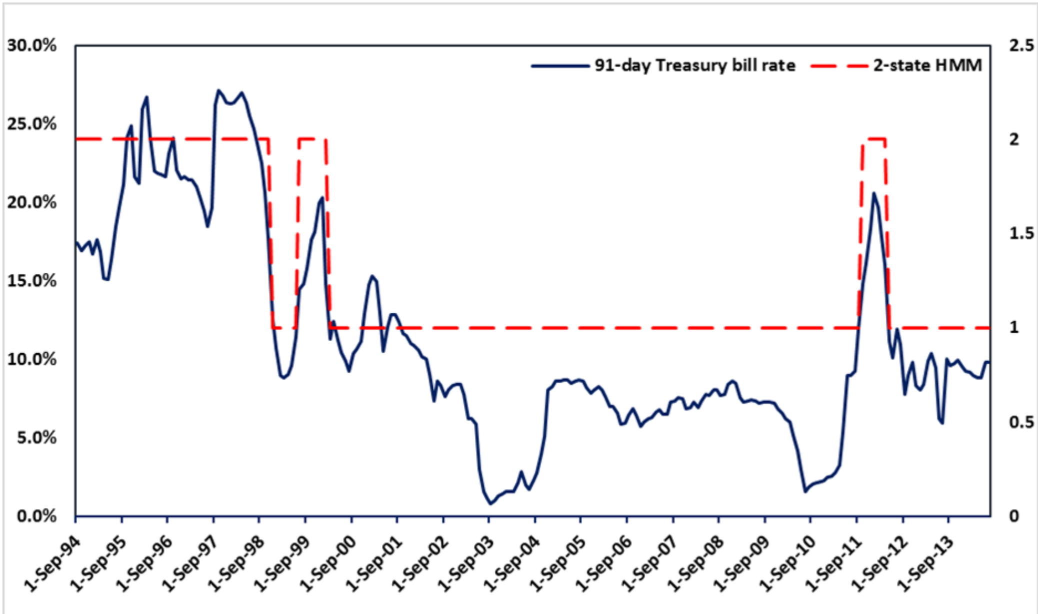

Figure 4.4: Evolution of the 2-state HMM model applied to the 91-day Treasury bill rate from September 1994 to July 2014

4.2.2 3-regime model

The HMM fitting procedure for a 3-state model converges at at 46 iterations with a LogLikelihood of 501.2552 and 14 degrees of freedom. Information criteria are obtained as -974.5105 for the AIC and -925.84 for the BIC.

The 3-state model, starts with the assumption that the interest rate process starts in state 2. Based on the HMM filtering, the transition probability matrix in figure 4.5 below is obtained:

Transition matrix	To State 1	To State 2	To State 3
From State 1	0.992	0.008	0.000
From State 2	0.032	0.951	0.016
From State 3	0.000	0.027	0.973

Figure 4.5: Transition probability matrix for the 3-state HMM model

The 3 states are observed to have a high same-state transition probability with very low other-state transition probabilities. It is interesting to note that there are no jumps from the extreme states (1 and 3) between each other. The transition probability matrix suggests that the movement of interest rates follows a gradual process, moving from one state, to the intermediate state and finally getting to its final state. The absence of sudden, extreme jumps suggest more stability for investors in interest rate dependent instruments.

By assuming the interest rate process is Gaussian, the descriptive statistics observed are shown in figure 4.6 below:

Response parameters	**Re 1. (Intercept)**	**Re 1. Sd**
State 1	0.066	0.026
State 2	0.138	0.033
State 3	0.232	0.026

Figure 4.6: Descriptive statistics for the 3-state HMM model

The graphical results of state changes for the 3-state HMM model are shown in figure 4.7 below:

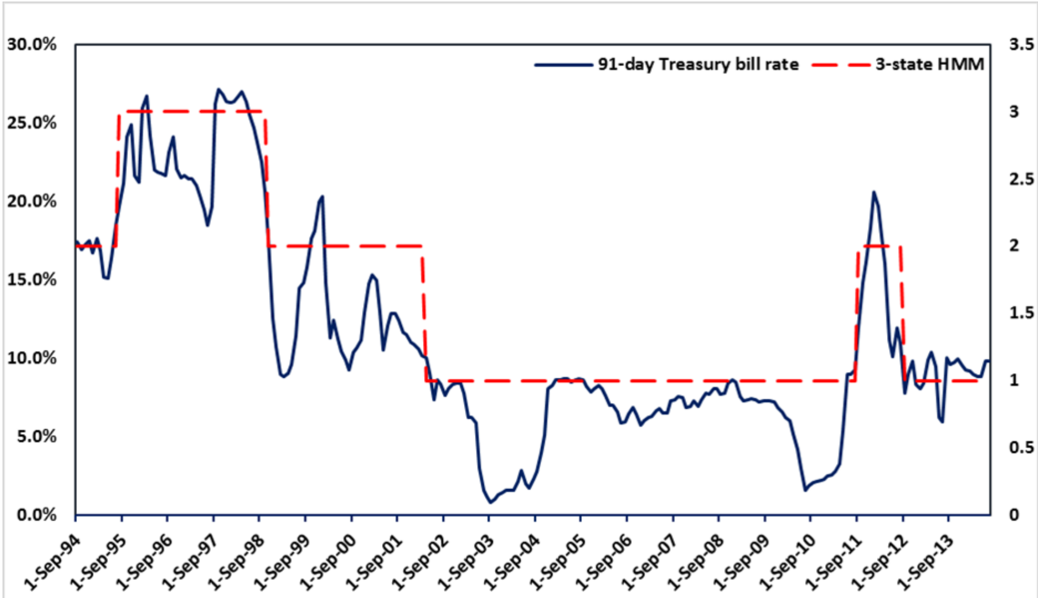

Figure 4.7: Evolution of the 3-state HMM model applied to the 91-day Treasury bill rate from September 1994 to July 2014

4.2.3 4-regime model

The HMM fitting procedure for a 4-state model converges at 34 iterations with a LogLikelihood of 597.2597 and 23 degrees of freedom. Information criteria are obtained as -1148.519 for the AIC and -1068.561 for the BIC.

The 4-state model, starts with the assumption that the interest rate process starts in state 3. Based on the HMM fitting, the transition probability matrix in figure 4.8 below is obtained:

Transition matrix	To State 1	To State 2	To State 3	To State 4
From State 1	0.930	0.070	0.000	0.000
From State 2	0.019	0.969	0.012	0.000
From State 3	0.000	0.036	0.947	0.017
From State 4	0.000	0.000	0.027	0.973

Figure 4.8: Transition probability matrix for the 4-state HMM model

As with the 3-state HMM model, the 4-state model also shows high same-state transition probability but with increased probability of transitioning to other states. State 3 for example has the probability of remaining in state 3 at 94.7 percent with transition to state 2 at 3.6 percent and the probability of transition to state 4 at 1.7 percent. As with the 3-state model, we also observe that there is no transition between states that are not immediately adjacent. For example, at no point does the interest rate move from state 3 to state 1.

By assuming the interest rate process is Gaussian, the descriptive statistics observed are shown in figure 4.6 below:

Response parameters	**Re1. (Intercept)**	**Re 1. Sd**
State 1	0.022	0.008
State 2	0.078	0.012
State 3	0.138	0.033
State 4	0.232	0.026

Figure 4.9: Descriptive statistics for the 4-state HMM model

The graphical results of state changes for the 4-state HMM model are shown in figure 4.10 below:

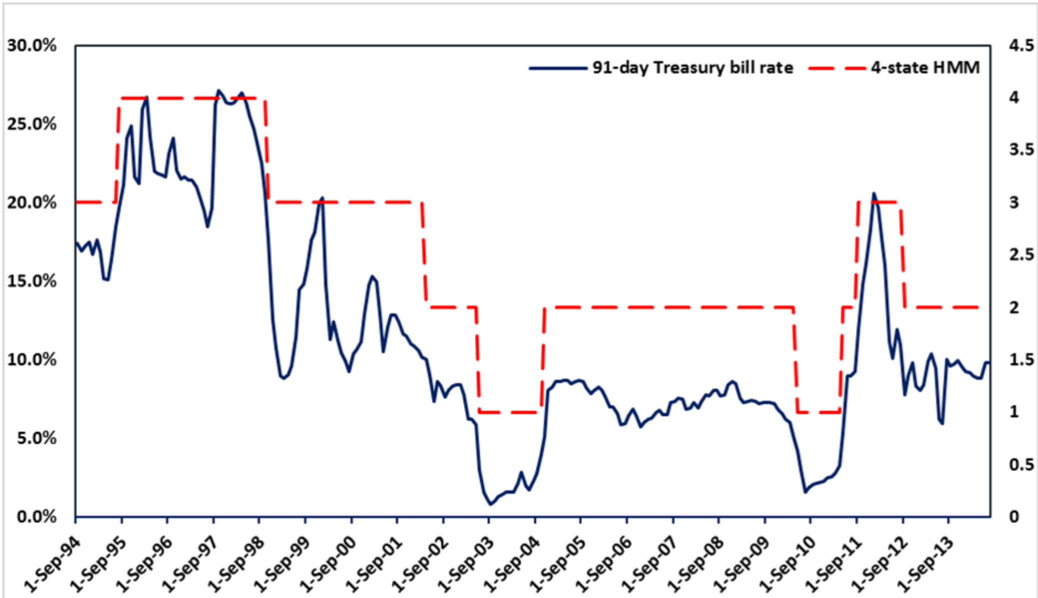

Figure 4.10: Evolution of the 4-state HMM model applied to the 91-day Treasury bill rate from September 1994 to July 2014

Chapter 5

Findings and conclusions

This section considers two key focus areas; 1)The results of HMM filtering to determine regimes, and 2) Calibration results for the Vasicek model.

5.1 Results of regime analysis

From the HMM results, different models are observed to have distinct representations for the interest rate process. We observe the 2-state model having an AIC of -877.0902 while the AIC for the 3-state model is -974.5105. This suggests better fit for the 3-state model compared to the 2-state model. Similarly, the AIC for the 4-state model is -1148.519 suggesting an improved fit compared to the 3-state model. This supports the hypothesis that the use of regime-switching models improves the fit of the model.

The results of the HMM regime breakdown are shown in figure 5.1 below:

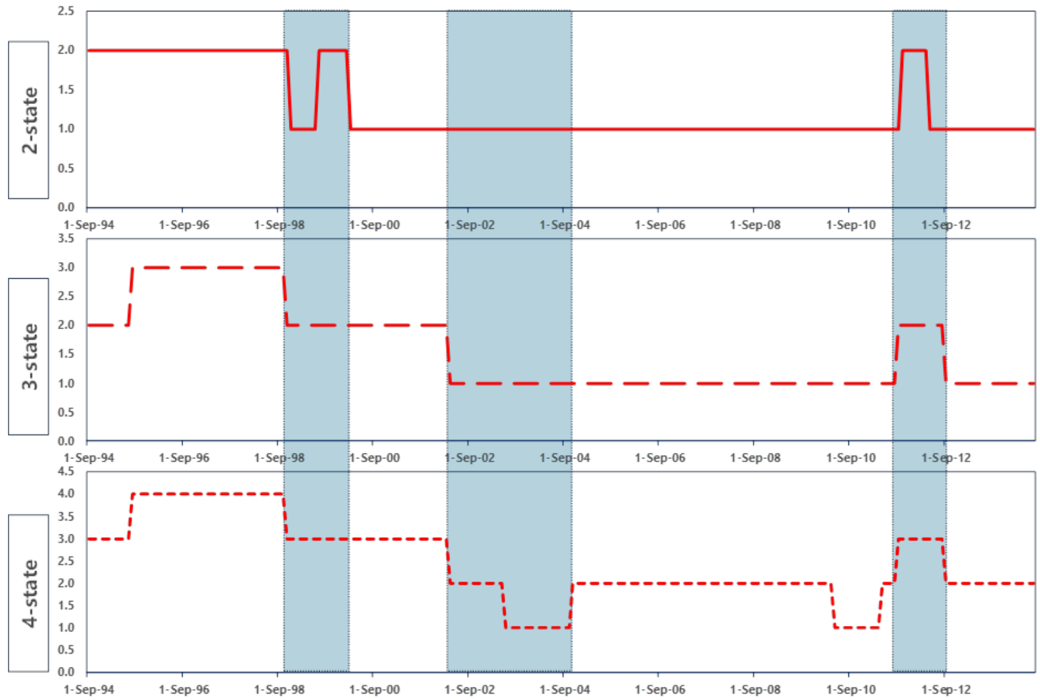

Figure 5.1: Comparison of state changes for the 2-state, 3-state and 4-state regime switching models applied to the 91-day Treasury bill rate from September 1994 to July 2014

From the above, it can be noted that:

- The two-state HMM model allows for a high average interest rate regime and a low average interest rate regime without additional differentiation. The low interest rate regime covers the larger proportion (72 percent) of the total sample period.
- The three-state HMM model provides for a high average interest rate regime and a low average interest rate regime with an additional transition regime that provides the bridge between the high and low average interest rate regimes.
- The four-state HMM model allows for a very-high interest rate state, a very low interest rate state, and two intermediate states representing both high and low interest rate levels.

5.2 Vasicek model parameter calibration

In calibrating the Vasicek model, the long-term mean (μ), the reversion rate (α), and the volatility (σ) are estimated. Maximum Likelihood estimation is used to determine the parameters per the equation:

$$dr(t) = \kappa[\mu - r(t)]dt + \sigma dW(t), r(0) = r_0 \tag{5.1}$$

where r_0, κ, and σ are positive constants.

5.2.1 Single-regime Vasicek model

For the single-regime Vasicek model, the long-term mean is obtained as 9.7773 percent. The rate of mean reversion (κ) is 2.1823 percent while the volatility is 1.3464 percent. The results are indicated in figure 5.2 below:

Table 5.1: Parameter estimates for the single-state Vasicek model for the 91-day Treasury bill rates for the period September 1994 to July 2014

Single-state model	Parameter estimates
The Long-term mean (μ)	9.773%
The mean-reversion rate (κ)	2.1823%
Instantaneous volatility (σ)	1.3464%
Asymptotic variance	0.4154%
Half-life ($t_{\frac{1}{2}}$)	13.79
Total (n)	239

The table shows a fairly high long term mean rate for the entire period under study. However, the observed rate is higher than that observed in the second half of the dataset, indicating that the parameter estimates do not show the changing dynamics of the interest rate process.

Given the observed mean-reversion rate (κ) of 2.1823 percent, we calculate the half-life of the mean-reversion process defined as $t_{\frac{1}{2}} = \frac{log2}{\kappa}$. For the single-state model, our estimate of the half life, $t_{\frac{1}{2}}$ is 13.79.

5.2.2 Multiple-regime Vasicek model

The Vasicek model parameters in different regimes are estimated based on results from the HMM filtering process. The results presented below show the different parameter estimates based on calibration of the Vasicek model under the different regimes.

2-state model

The 2-state Vasicek model breaks the interest rate process into two states with distinct parameter estimates. State 1 is observed to have a low long-term mean and low volatility compared to State 2. From the tables below, we see the breakdown of the interest rate process into different regimes based on the long-term mean.

Table 5.2: Parameter estimates for the 2-state Vasicek model for the 91-day Treasury bill rates from September 1994 to July 2014

2-state model	State 1	State 2
The Long-term mean (μ)	6.8923%	21.4387%
The mean-reversion rate (κ)	11.0408%	15.7224%
Instantaneous volatility (σ)	1.0716%	1.7016%
Asymptotic variance	0.0520%	0.0921%
Half-life ($t_{\frac{1}{2}}$)	2.73	1.91
Total (n)	173	66

The table shows the parameter estimates for the 2-state model which breaks out the interest rate series into a high interest rate and a low-interest rate series. There is only marginal differentiation in the instantaneous volatility but with a noticeable increase in the mean-reversion parameters for the two states. These results show higher stability of the different states with clear separation between the high and low long-term mean rates.

As with the single-state model, we calculate the half-life of the different states. We observe the half-life as 2.73 for state 1 and 1.91 for state 2. This results shows that the 2-state model takes a shorter time to be pulled back to the mean compared to the single-state model with a half life of 13.79.

3-state model

Table 5.3: Parameter estimates for the 3-state Vasicek model for the 91-day Treasury bill rates from September 1994 to July 2014

3-state model	State 1	State 2	State 3
The Long-term mean (μ)	6.3750%	13.1614%	23.4144%
The mean-reversion rate (κ)	6.8761%	17.1562%	25.6329%
Instantaneous volatility (σ)	0.8828%	1.7125%	1.6224%
Asymptotic variance	0.0567%	0.0855%	0.0513%
Half-life ($t_{\frac{1}{2}}$)	4.38	1.75	1.17
Total (n)	136	64	39

The parameter estimates under the 3-state model show the breakdown into three states: a high interest rate state, a low-interest rate state, and a transitional state with the long-term mean parameter estimates reflecting this result. The instantaneous volatility estimates follow this pattern with state 1 having a low volatility, state 2 having a higher volatility but with the highest volatility observed for the transitional state, state 2.

Given the estimated parameters, we compute the half-life for the different states and obtain: 4.38 for state 1, 1.75 for state 2 and 1.17 for state 3. All states in the 3-state model show a much lower half-life compared to the single-regime model. Similarly, states 2 and 3 show lower half-lives compared to the 2-state model with state 1 having a slightly higher reversion rate.

4-state model

Table 5.4: Parameter estimates for the 4-state Vasicek model for the 91-day Treasury bill rates from September 1994 to July 2014

4-state model	State 1	State 2	State 3	State 4
The Long-term mean (μ)	2.0404%	7.8077%	13.1614%	23.4144%
The mean-reversion rate (α)	49.4913%	32.7593%	17.1562%	25.6329%
Instantaneous volatility (σ)	0.5755%	0.8131%	1.7125%	1.6224%
Asymptotic variance	0.0033%	0.0101%	0.0855%	0.0513%
Half-life ($t_{\frac{1}{2}}$)	0.61	0.92	1.75	1.17
Total (n)	29	107	64	39

The table showing calibration results for the 4-state model shows 2 high-interest rate, and 2 low-interest rate regimes. This pattern is also observed for the instantaneous volatility.

The 4-state model shows the lowest values for half life across all states. The half lives are obtained as 0.61 for state 1, 0.92 for state 2, 1.75 for state 3 and 1.17 for state 4. From the analysis, it is observed that the 4-state model breaks down state 1 in the 3-state model into two separate states both having very low half-lives.

5.2.3 Evaluation of alternative models

The HMM filtering approach provides a clear separation of states relying significantly on the level of the interest rate. The level of the long-term mean is a significant determinant of the observed state. This is the case for the 2-state, 3-state and 4-state models.

By contrast, the volatility, and by extension the asymptotic variance does not follow a similar pattern. There is no direct observed relationship between the level of the interest rate process and the observed volatility. This suggests that the risk and expected compensation does not depend on the level of the interest rate but rather the inherent variability of the interest rate while at a particular interest rate level.

The mean reversion rate shows the most significant variation with change for different states and across the different HMM models. The mean reversion rate is 2.1823 percent for the single state model but the weighted average mean reversion rate jumps to 12.3336 percent for the 2-state model. This increases to 12.6897 percent for the 3-state model and jumps to 29.4484 percent for the 4-state model. This observation indicates that the addition of states increases the mean-reversion parameter for the individual states. This suggests that the addition of states improve stability of the long-term mean in each state. This makes intuitive sense as a higher mean reversion rate implies faster reversion to the long-term mean level. This is supported by calculations of the half-lives for the different states which show a steady decrease with the introduction of additional states.

5.3 Conclusion

In this paper, the Vasicek model has been calibrated under a single regime and under multiple regimes of the interest rate process. It is observed that introduction of regimes enables more flexible specification of the interest rate process. Three key observations are made:

- The level of the long-term mean is a significant determinant of the state probability.
- The volatility and asymptotic variance do not depend on the level of the long-term mean, but on the inherent variability of the process in a particular state.
- The introduction of additional state probabilities increases the mean reversion parameter for majority of the states. This suggests increased stability in model parameters given the introduction of regimes.

It is observed that a regime-switching model of the interest rate provides a good framework for modeling the interest rate and as such should be considered when evaluating interest rate dynamics for the pricing of derivative instruments. The parameter estimates provide a good starting point for development of financial instruments whose underlying is or depends on the interest rate process.

Future areas of research include evaluating the comparative performance of the Vasicek interest rate model against other interest rate models. In addition, there would be significant value in evaluating the evolution of interest rate dynamics and volatility when constraints are applied to the interest rate process e.g., with politically-motivated interest rate controls. Additionally, this work focused on the calibration with scope for future work to evaluate the forecasting results.

Existing models used by financial institutions in Kenya focus on economic factors e.g., economic growth, inflation etc, which present an opportunity to evaluate the performance of diffusion models when compared to models based on underlying economic phenomena. This analysis can also be extended to evaluate the impact of regime-switches on credit risk measures. These are areas that would provide additional context on the most appropriate models for empirical application in a developing market like Kenya.

References

A. Ang and G. Bekaert, "Regime switches in interest rates," *National Bureau of Economic Research*, 2001.

T. Bjork, *Arbitrage Theory in Continuous time*, 3rd ed. Oxford University Press, 2009.

D. Brigo, "Interest rate models: Paradigm shifts in recent years," 2007.

D. Brigo and F. Mercurio, *Interest Rate Models: Theory and Practice (with Smile, Inflation and Credit).* Springer Verlag, 2006.

D. Brigo, A. Dalessandro, M. Neugebauer, and F. Triki, "A stochastic processes toolkit for risk management. mean reverting processes and jumps." *Journal of Risk management in Financial institutions*, 2008.

G. M. Caporale and L. A. Gil-Alana, "Interest rate dynamics in Kenya: Commercial banks' rates and the 91-day treasury bill rate." *Brunel University, Economics and Finance Working Paper Series, Working paper no. 10-27*, 2010.

J. Cox, J. E. Ingersoll, and S. A. Ross, "A theory of the term structure of interest rates," *Econometrica*, vol. 53, no. 2, pp. 385–408, 1985.

C. Cuchiero, "Affine interest rate models - theory and practice," Master's thesis, Vienna: Technische Universitat Wien - Vienna Univeristy of Technology., 2006.

M. Dahlquist and F. S. Gray, "Regime-switching and interest rates in the European monetary system," *Journal of International Economics*, 1998.

R. Elliott and C. Wilson, "The term structure of interest rates in a hidden Markov setting," 2007.

C. Erlwein, “Applications of hidden Markov models in financial modeling,” Ph.D. dissertation, Brunel University, 2008.

F. S. Gray, “Modeling the conditional distribution of interest rates as a regime-switching process,” *Journal of Financial Economics*, 1996.

D. Heath, R. Jarrow, and A. Morton, “Bond pricing and the term structure of interest rates: A new methodology for contingent claims valuation,” *Econometrica*, vol. 60, no. 1, pp. 77–105, January 1992.

J. Maheu and Q. Yang, “An infinite hidden markov model for short-term interest rates,” Ph.D. dissertation, McMaster University and University of Toronto, 2015.

R. Munclinger, “Exploration of the Brazilian term structure in a hidden Markov framework,” International Monetary Fund (IMF), Tech. Rep., 2011.

M. Musiela and M. Rutkowski, *Martingale Methods in Financial Modeling*, 2nd ed., 2005.

W. Ngugi, R., “An empirical analysis of interest rate spread in Kenya,” *Africa Economic Research Consortium*, vol. Research paper 106, 2001.

T. Olweny, “Modeling volatility of short-term interest rates in Kenya,” *International Journal of Business and Social Science*, pp. 289–303, 2011.

S. Svoboda, “An investigation of various interest rate models and their calibration in the South African market.” Ph.D. dissertation, University of Witwatersrand, Johannesburg, 2002.

E. Van Elen, “Term-structure forecasting: Does a good fit imply reasonable simulation results?” *Netspar Theses*, 2010.

O. Vasicek, “An equilibrium characterization of the term structure,” *Journal of Financial Economics*, vol. 5, pp. 177–188, August 1977.

S. Wu and Y. Zeng, “Affine regime-switching models for interest rate term structure,” *Journal of Contemporary Mathematics*, 1991.