Earlier title: Estimation of some simple diffusion models for financial data

Earlier title: Estimation of some simple diffusion models for financial data

Earlier title: Statistical analysis of interest rate models when trading is infrequent

Earlier title: Estimation of some simple diffusion models for financial data

Earlier title: Statistical analysis of interest rate models when trading is infrequent

Helgi Tómasson

mailto:helgito@hi.is

8th November 2005

Seminar Central Bank of Iceland

- **1** Organization
 - Background

- Background
- Taylor expansion of Kolmogorov forward equation

- Background
- Taylor expansion of Kolmogorov forward equation
- Description of calculation of approximations

- Background
- Taylor expansion of Kolmogorov forward equation
- Description of calculation of approximations
- Simulation of a univariate model

- Background
- Taylor expansion of Kolmogorov forward equation
- Description of calculation of approximations
- Simulation of a univariate model
- Simulation of a multivariate model

- Background
- Taylor expansion of Kolmogorov forward equation
- Description of calculation of approximations
- Simulation of a univariate model
- Simulation of a multivariate model
- A simple application to Icelandic bond data
- Discussion

• Financial data, tick-by-tick

- Financial data, tick-by-tick
- Estimation of diffusion models

- Financial data, tick-by-tick
- Estimation of diffusion models
- Approximations of the transition density for diffusion

- Financial data, tick-by-tick
- Estimation of diffusion models
- Approximations of the transition density for diffusion
- Taylor expansion of Kolmogorov equations

- Financial data, tick-by-tick
- Estimation of diffusion models
- Approximations of the transition density for diffusion
- Taylor expansion of Kolmogorov equations

Seminar Central Bank of Iceland

• What is a diffusion model?

- What is a diffusion model?
- Answer: A stochastic model describing continuous time dynamics. Such models have been made popular in mathematical fincance, e.g., by Nobel-prize winners Merton and Scholes.

- What is a diffusion model?
- Answer: A stochastic model describing continuous time dynamics. Such models have been made popular in mathematical fincance, e.g., by Nobel-prize winners Merton and Scholes.
- A typical way of representing such models, i.e. desribing the nature of a dynamic process X(t), is by means of stochastic differential equations:

$$dX(t) = \underbrace{\mu(X(t), \boldsymbol{\theta})dt}_{\text{drift}} + \underbrace{\sigma(X(t), \boldsymbol{\theta})dW(t)}_{\text{diffusion term}}$$

 $\boldsymbol{\theta}$ is að vector of parameters that define the behaviour of the process.

Seminar Central Bank of Iceland

- W(t) is a standard Wiener process. For t > s, E(W(t)|W(s)) = W(s), W continuous independent increment process. dW(t) is continuous time white noise.
- The $\mu(X(t), \theta)dt$ part represents the predictable part of the process.
- The $\sigma(X(t), \theta) dW$ part represents the stochastic part.

Data and model

• The mathematical idealization

Data and model

- The mathematical idealization
- A given diffusion:

$$dX(t) = \mu(X(t), \boldsymbol{\theta})dt + \sigma(X(t), \boldsymbol{\theta})dW(t)$$

is observed at times t_1, \ldots, t_n . The parameter, θ is to be estimated from $X(t_1), \ldots, X(t_n)$.

Data and model

- The mathematical idealization
- A given diffusion:

$$dX(t) = \mu(X(t), \theta)dt + \sigma(X(t), \theta)dW(t)$$

is observed at times t_1, \ldots, t_n . The parameter, θ is to be estimated from $X(t_1), \ldots, X(t_n)$.

 A few approaches, simulation methods, method of moments, estimating functions, maximum-likelihood approximation. • The transition density:

 $f(x|x_0, \Delta) =$ density for $X(t + \Delta)$ given $X(t) = x_0$

is only known for for some specific $\mu(X(t), \pmb{\theta})$ and $\sigma(X(t), \pmb{\theta})$

Therefore maximizing the log-likelihood, i.e., solving:

$$\max_{\boldsymbol{\theta}} l(\boldsymbol{\theta}|X(t_1),\ldots,X(t_n))$$

is only possible in some special cases.

Some populuar diffusion models

OU
$$dX(t) = \kappa(\alpha - X(t))dt + \sigma dW(t)$$
 (1)
Ornstein-Uhlenbeck/Vasicek
CIR $dX(t) = \kappa(\alpha - X(t))dt + \sigma\sqrt{X(t)}dW(t)$ (2)
Cox-Ingersoll-Ross/square-root process
CKLS $dX(t) = \kappa(\alpha - X(t))dt + \sigma X(t)^{\rho}dW(t)$ (3)
Chan, Karolyi, Longstaff & Sanders (1992),

Cases of special interest $\rho = 1/2$ and $\rho = 1$

These are all stochastic versions of a very simple differential equation:

$$dX(t) = \kappa(\alpha - X(t))dt$$

Given X(0), the solution is of the form:

$$X(t) = \alpha + \exp(-\kappa t)(X(0) - \alpha)$$

Seminar Central Bank of Iceland

The α parameter is the longtime equilibrium, κ controls the speed of convergence to equilibrium.

The α parameter is the longtime equilibrium, κ controls the speed of convergence to equilibrium.

If at time 0, the system is at $(X(0) - \alpha)$ distance from equilibrium, then it will take the system $\log(2)/\kappa$ time units to decrease that distance by 50%.

- The α parameter is the longtime equilibrium, κ controls the speed of convergence to equilibrium.
- If at time 0, the system is at $(X(0) \alpha)$ distance from equilibrium, then it will take the system $\log(2)/\kappa$ time units to decrease that distance by 50%.
- In the stochastic case σ and ρ control the volality of the process.

Kolmogorov forward equation

• What is known about $f(x|x_0, \Delta)$?, $x = x(t + \Delta)$, $x_0 = x(t)$

Kolmogorov forward equation

- What is known about $f(x|x_0, \Delta)$?, $x = x(t + \Delta)$, $x_0 = x(t)$
- Since X(t) is a diffusion process the density function $f(x|x_0, \Delta)$ solves:

$$\frac{\partial f(x|x_0,\Delta)}{\partial \Delta} + \frac{\partial \left(\mu(x,\boldsymbol{\theta})f(x|x_0,\Delta)\right)}{\partial x} \\ -\frac{1}{2}\frac{\partial^2 \left(\sigma^2(x,\boldsymbol{\theta})f(x|x_0,\Delta)\right)}{\partial x^2} = 0$$

What can be said about it?

Assuming $\sigma(x) = 1$,

Assuming $\sigma(x) = 1$,

writing $p(x|x_0, \Delta) = log(f(x|x_0, \Delta))$

- Assuming $\sigma(x) = 1$,
- writing $p(x|x_0, \Delta) = log(f(x|x_0, \Delta))$
- and substituting $e^{p(x|x_0,\Delta)}$ for $f(x|x_0,\Delta)$ in Kolmogorov's equations gives:

Assuming $\sigma(x) = 1$,

writing $p(x|x_0, \Delta) = log(f(x|x_0, \Delta))$

and substituting $e^{p(x|x_0,\Delta)}$ for $f(x|x_0,\Delta)$ in Kolmogorov's equations gives:

$$\frac{\partial p(x|x_0,\Delta)}{\partial \Delta} + \mu'(x) + \mu(x)\frac{\partial p(x|x_0,\Delta)}{\partial x}$$
$$-\frac{1}{2}\left[\frac{\partial p(x|x_0,\Delta)}{\partial x}\right]^2 - \frac{1}{2}\frac{\partial^2 p(x|x_0,\Delta)}{\partial x^2} = 0$$

(4)

A Taylor expansion of $p(x|x_0, \Delta)$ in Δ is on the form:
A Taylor expansion of $p(x|x_0, \Delta)$ in Δ is on the form:

$$-\frac{1}{2}log(2\pi\Delta) - \frac{(x-x_0)^2}{2\Delta} + c_0(x|x_0) + c_1(x|x_0)\Delta$$
$$+c_2(x|x_0)\frac{\Delta^2}{2} + c_3(x|x_0)\frac{\Delta^3}{3!} + \cdots$$

Substituting the Taylor expansion into equation (4) gives the first two terms:

Substituting the Taylor expansion into equation (4) gives the first two terms:

$$-\frac{(x-x_0)(\mu(x) - c'_0(x|x_0))}{\Delta}$$
(5)
$$-\frac{1}{2}c_0(x|x_0)'(x)^2 + \mu(x) + \mu(x)c'_0(x|x_0)$$
(6)
$$-\frac{c''_0(x|x_0)}{2} + c_1(x|x_0) + (x-x_0)c'_1(x|x_0)$$

Next terms

Next terms

$$+\frac{1}{2}\Delta(2(\mu(x) - c'_0(x|x_0))c'_1(x|x_0) - c''_1(x|x_0)) + 2c_2(x|x_0) + (x - x_0)c'_2(x|x_0))$$

$$+\frac{1}{12}\Delta^{2}(-6c_{1}'(x|x_{0})^{2}+6(\mu(x)-c_{0}'(x|x_{0}))c_{2}'(x|x_{0}))\\-3c_{2}''(x|x_{0})+6c_{3}(x|x_{0})+2(x-x_{0})c_{3}'(x|x_{0}))$$

- •
- •
- -

The Kolomogorov equations force the coefficients for each power of Δ to be zero.

Equation (5) gives

$$c_0(x|x_0)=\int_{x_0}^x \mu(u)du$$

• Substituting $c_0(x|x_0)$ into reduces the system of equations to

$$\frac{1}{2}(\mu(x)^2 + \mu'(x)^2) + c_1(x|x_0) + (x - x_0)c'_1(x|x_0) = 0$$

- $c''_1(x|x_0) + 2c_2(x|x_0) + (x - x_0)c'_2(x|x_0) = 0$
- $\frac{3}{2}c''_2(x|x_0) - 3c'_1(x|x_0)^2$
+ $3c_3(x|x_0) + (x - x_0)c'_3(x|x_0) = 0$

•

And more

And more

$$-2c_{3}''(x|x_{0}) - 12c_{1}'(x)c_{2}'(x) +$$

$$4c_{4}(x|x_{0}) + (x - x_{0})c_{4}'(x|x_{0}) = 0$$

$$-\frac{5}{2}c_{4}''(x|x_{0}) - 20c_{1}'(x|x_{0})c_{3}'(x|x_{0})$$

$$-15c_{2}'(x|x_{0})^{2} + 5c_{5}(x|x_{0}) + (x - x_{0})c_{5}'(x|x_{0}) = 0$$

$$\vdots$$

Seminar Central Bank of Iceland

 The c_j functions are derived recursively by solving the differential equations of the type:

$$jc_j(x|x_0) + (x - x_0)c'_j(x|x_0) = q_j(x) \quad \text{which gives}$$
$$c_j(x) = \frac{1}{(x - x_0)^j} \int_{x_0}^x (u - x_0)^{j-1} q_j(u) du$$

• The functions $q_j(x)$ are deciced by c_0, \ldots, c_{j-1}

Some comments

• What if the diffusion is not 1?

Some comments

- What if the diffusion is not 1?
- Transform X(t) to Y(t) such that Y(t) has diffusion 1. That is always possible if X(t) is one-dimensional

Some comments

- What if the diffusion is not 1?
- Transform X(t) to Y(t) such that Y(t) has diffusion 1. That is always possible if X(t) is one-dimensional

$$Y(t) = \gamma(X(t)) = \pm \int_{0}^{X(t)} \frac{du}{\sigma(u)} + \text{constant}$$

• Ito's lemma gives that Y(t) will have unit diffusion and drift:

$$\mu_Y(y) = \pm \left(\frac{\mu(\gamma^{-1}(y))}{\sigma(\gamma^{-1}(y))} - \frac{1}{2}\frac{\partial\sigma}{\partial x}(\gamma^{-1}(y))\right)$$

• The densities f_X and f_Y are related by:

 $f_X(x|x_0, \Delta) = f_Y(y|y_0, \Delta) |\text{Jacobian}| = f_Y(y|y_0, \Delta) / \sigma(x)$

• The densities f_X and f_Y are related by:

 $f_X(x|x_0, \Delta) = f_Y(y|y_0, \Delta) |\text{Jacobian}| = f_Y(y|y_0, \Delta) / \sigma(x)$

• I.e. the connection between the log-densites for X(t)|X(0) and Y(t)|Y(0) is:

$$p_X(x|x_0, \Delta) = p_Y(y|y_0, \Delta) + \log(1/\sigma(x))$$

• Therefore the transformation is only a technicality. I.e. the Taylor coefficients c_j will be messier functions of x and x_0 than of $y = \gamma(x)$ and $y_0 = \gamma(x_0)$.

- Therefore the transformation is only a technicality. I.e. the Taylor coefficients c_j will be messier functions of x and x_0 than of $y = \gamma(x)$ and $y_0 = \gamma(x_0)$.
- Numerically it might be better to work with Y(t) than X(t).

- Therefore the transformation is only a technicality. I.e. the Taylor coefficients c_j will be messier functions of x and x_0 than of $y = \gamma(x)$ and $y_0 = \gamma(x_0)$.
- Numerically it might be better to work with Y(t) than X(t).
- Solving recursively for the functions c_j becomes increasingly complicated. A Taylor expansion in xaround x_0 (or y around y_0) is therefore an option.

 If no transformation took place, one still gets a sequence of differential equations to solve, but they will be more complicated.

$$2c_{-1}(x) + \sigma(x)^2 (c'_{-1}(x))^2 = 0 \quad c_{-1}(x_0) = 0$$

will be the first one

Recursive system for untransformed variable:

Recursive system for untransformed variable:

Condition (1) $v(x) = \sigma^2(x)$

$$v(x)c'_{-1}(x)^2 + 2c_{-1}(x) = 0$$

$$c_{-1}(x) = -\frac{1}{2} \left(\int_{x_0}^x \frac{du}{\sqrt{v(x)}} \right)^2$$

Condition(2)

$$-c_0'(x)c_{-1}'(x)v(x)^2 - \frac{1}{2}c_{-1}''(x)v(x)^2$$
$$-\frac{3}{2}c_{-1}'(x)v'(x)v(x) + \mu(x)c_{-1}'(x) - \frac{1}{2} = 0$$

Condition (3) (Coefficient on t)

$$c_{1}(x) - c'_{1}(x)c'_{1}(x)\sigma(x)^{2} + \mu(x)c'_{0}(x) + \mu'(x) - \frac{\mu(x)\sigma'(x)}{2\sigma(x)} - \frac{1}{2}c'_{0}(x)^{2}\sigma(x)^{2} - \frac{1}{2}c''_{0}(x)\sigma(x)^{2} - \frac{3}{2}c'_{0}(x)\sigma'(x)\sigma(x) - \frac{3}{4}\sigma''(x)\sigma(x) - \frac{3}{8}\sigma'(x)^{2} = 0$$

Higher dimensions

• Same principles apply

Higher dimensions

- Same principles apply
- Think of a 2-dimensional case.

$$d\begin{bmatrix} X_1(t) \\ X_2(t) \end{bmatrix} = \begin{bmatrix} \mu_1(X_1(t), X_2(t), \boldsymbol{\theta}) \\ \mu_2(X_1(t), X_2(t), \boldsymbol{\theta}) \end{bmatrix} dt + \begin{bmatrix} \sigma_{11}(X_1(t), X_2(t), \boldsymbol{\theta}) & \sigma_{12}(X_1(t), X_2(t), \boldsymbol{\theta}) \\ \sigma_{21}(X_1(t), X_2(t), \boldsymbol{\theta}) & \sigma_{21}(X_1(t), X_2(t), \boldsymbol{\theta}) \end{bmatrix} \begin{bmatrix} dW_1(t) \\ dW_2(t) \end{bmatrix}$$

• The log-density is assumed to be of the form

$$-\frac{m}{2}log(2\pi\Delta) - D(\boldsymbol{x},\boldsymbol{\theta}) + \frac{C^{(-1)}(\boldsymbol{x},\boldsymbol{\theta})}{\Delta} + C^{(0)}(\boldsymbol{x},\boldsymbol{\theta}) + C^{(1)}(\boldsymbol{x},\boldsymbol{\theta})\Delta + C^{(2)}(\boldsymbol{x},\boldsymbol{\theta})\Delta^2/2 + C^{(3)}(\boldsymbol{x},\boldsymbol{\theta})\Delta^3/3! + \cdots$$
$$D(\boldsymbol{x},\boldsymbol{\theta}) = \frac{1}{2}log(det(\sigma(\boldsymbol{x},\boldsymbol{\theta})\sigma(\boldsymbol{x},\boldsymbol{\theta})^T)$$

• Some conditions on σ_{ij} are needed in order to find a neat transformation as in the univariate case.

- Some conditions on σ_{ij} are needed in order to find a neat transformation as in the univariate case.
- Ait-Sahalia calls that situation a ,,reducible'' case, and the case where such a transformation does not exist.

• Analytical solutions of the $C^{(j)}$ functions difficult, so focus is on Taylor expansions (in x).

- Analytical solutions of the $C^{(j)}$ functions difficult, so focus is on Taylor expansions (in x).
- The resulting functions will be C^(j,l) where an *l*-th order Taylor approximation of order *l* has been taken in *x*. A trick that could also be useful in one dimension.

- Analytical solutions of the $C^{(j)}$ functions difficult, so focus is on Taylor expansions (in x).
- The resulting functions will be C^(j,l) where an *l*-th order Taylor approximation of order *l* has been taken in *x*. A trick that could also be useful in one dimension.
- The C^(j,l) functions can be derived analogously to the univariate case in a messy but straightforeward manner.

A small simulation. The model

$$dX = \kappa(\alpha - X)dt + \sigma X^{\rho}dW$$

A small simulation. The model

$$dX = \kappa(\alpha - X)dt + \sigma X^{\rho}dW$$

is simulated using Milstein-scheme (strong Taylor of order 1, 25 replications).

• The time spans used are T = 1, 10, 100

A small simulation. The model

$$dX = \kappa(\alpha - X)dt + \sigma X^{\rho}dW$$

- The time spans used are T = 1, 10, 100
- $\Delta = 1$, 0.1 and 0.01 are used.

A small simulation. The model

$$dX = \kappa(\alpha - X)dt + \sigma X^{\rho}dW$$

- The time spans used are T = 1, 10, 100
- $\Delta = 1$, 0.1 and 0.01 are used.
- 10 points of process per observation.

A small simulation. The model

$$dX = \kappa(\alpha - X)dt + \sigma X^{\rho}dW$$

- The time spans used are T = 1, 10, 100
- $\Delta = 1$, 0.1 and 0.01 are used.
- 10 points of process per observation.
- $\kappa = 0.24$, $\alpha = 0.07$, $\sigma = 0.08838$, $\rho = 0.75$.
| | T=1 | T=10 | T=100 |
|----------------|--------|--------|--------|
| $\hat{\kappa}$ | 3.2742 | 0.5352 | 0.2709 |
| \hat{lpha} | 0.0793 | 0.0962 | 0.0695 |
| $\hat{\sigma}$ | 0.1115 | 0.0979 | 0.0899 |
| $\hat{ ho}$ | 0.7695 | 0.7732 | 0.7570 |

Table 1: Average estimates, for $\Delta = 0.01$

	delta=1	delta=0.1
$\hat{\kappa}$	0.8232	0.2916
\hat{lpha}	0.0644	0.0744
$\hat{\sigma}$	0.0984	0.0864
$\hat{ ho}$	0.7342	0.7299

Table 2: Average estimates, for T=100

	T=1	T=10	T=100
s.d. $\hat{\kappa}$	1.8703	0.3995	0.0693
s.d $\hat{\alpha}$	0.0234	0.1269	0.0048
s.d. $\hat{\sigma}$	0.0542	0.0309	0.0042
s.d $\hat{ ho}$	0.2794	0.1084	0.0176

Table 3: Standard deviation of simulations, for Δ =0.01

	delta=1	delta=0.1	
s.d. $\hat{\kappa}$	0.4680	0.1807	
s.d \hat{lpha}	0.0099	0.0248	
s.d. $\hat{\sigma}$	0.0510	0.0196	
s.d $\hat{ ho}$	0.2190	0.0851	

Table 4: Standard deviation of simulations, for T=100

A two dimensional example:

$$dX = \mu dt + \sigma_1 e^Y dW_1$$
$$dY = \kappa (\alpha - Y) dt + \sigma_2 dW_2$$

Results of 20 replications, $\mu = 0$, $\sigma_1 = \sigma_2 = 1$, $\kappa = 20$, $\alpha = 0.01$. T=1, $\Delta = 1/10000$

	$\hat{\kappa}$	\hat{lpha}	$\hat{\sigma}_1$	$\hat{\mu}$	$\hat{\sigma}_2$
\bar{m}	24.1	0.0053	1.004	-0.159	1.000
sd	5.61	0.0519	0.006	0.983	0.007

Table 5: Simulation of 2-dim. model

Icelandic interest rate data

- Each transaction in 2002-2004.
- Zero-coupon governmental bonds, annualized to r(t).
- Form of data:
 RIKV 02 06 05 96.165 03.01.2002 11:42:4
 RIKV 02 08 06 94.915 04.01.2002 11:13:1

r(t) 2002-2004



Figure 1: 1000 days of Icelandic bond market

- Timeperiod is 1050 days. Trading took place on 432 days.
- 1933 transactions took place 5 days a week.

383 410 351 452 337



Figure 2: Trading frequency

Daily standard deviations of r(t)



Figure 3: Daily standard deviations

Daily range of r(t)



Figure 4: Daily observed range

- Of the 1933 observations, 664 had $\Delta_i = 0$. Many had very small $\Delta_i = 0$.
- The variance of the prices of the simultaneous observations offers a possibility to estimate the market microstructure noise. The standard deviation of simultaneous transactions is 6.4 points. (1%=100 points).
- For the expansion framework to work, the Δ_i 's have to be small, but not to small, e.g., 10^{-8} is to small. The diffusion models rule out large jumps in small intervals.

- 2 Result of CIR for Icelandic data
 - t=1 month. Median r(t) of a day chosen.

 $\hat{\kappa} = 15.090824$ $\hat{\alpha} = 0.046586$ $\hat{\sigma} = 0.081569$ s.e.=0.028743 s.e.=0.000007 s.e.=0.000237

Looks much more peaceful than the test example that Chan, Karolyi, Longstaff & Sanders (1992) claim is a natural result.

Test example

T=1000, $\Delta = 0.04$, simulation result (1000 time-periods), one replication. Parameter values from Chan, Karolyi, Longstaff & Sanders (1992).

<i>κ</i> =0.24	alpha = 0.08	$\sigma=$ 0.08838
$\hat{\kappa}=$ 0.23908	$\hat{\alpha} = 0.08353$	$\hat{\sigma} = 0.08757$
s.e.=0.02196	s.e.=0.00335	s.e.=0.00039

• $\bar{x} = 0.08358$, $\mu = \alpha = 0.08$

• s = 0.03556, max(x(t)) = 0.2801, min(x(t)) = 0.0089

• St.dev of stationary dist.= $\sqrt{\alpha * \sigma^2/(2 * \kappa)} = 0.03608$



Figure 5: A simulated CIR for 1000 time-periods.

 The Taylor approximations in the spirit of Ait-Sahalia (1999, 2002) seem to work well approximating maximum-likelihood estimates for simple diffusions.

- The Taylor approximations in the spirit of Ait-Sahalia (1999, 2002) seem to work well approximating maximum-likelihood estimates for simple diffusions.
- For exact estimates of the trend part of a diffusion a long span of data is required.

- The Taylor approximations in the spirit of Ait-Sahalia (1999, 2002) seem to work well approximating maximum-likelihood estimates for simple diffusions.
- For exact estimates of the trend part of a diffusion a long span of data is required.
- The parameters of a univariate diffusions might vary over time. It is of value for say, central banks to monitor, e.g., α and σ in a CIR-type model.

- The Taylor approximations in the spirit of Ait-Sahalia (1999, 2002) seem to work well approximating maximum-likelihood estimates for simple diffusions.
- For exact estimates of the trend part of a diffusion a long span of data is required.
- The parameters of a univariate diffusions might vary over time. It is of value for say, central banks to monitor, e.g., α and σ in a CIR-type model.

• The use of diffusion process for statistical analysis of dynamic processes is a feasible approach,

- The use of diffusion process for statistical analysis of dynamic processes is a feasible approach,
- In the authors mind more to bond data than very high frequency stock data but also in continuous-time macroeconometrics.

- The use of diffusion process for statistical analysis of dynamic processes is a feasible approach,
- In the authors mind more to bond data than very high frequency stock data but also in continuous-time macroeconometrics.

References

- Ait-Sahalia, Y. (1999). Transition densities for interest rate and other nonlinear diffusion. Journal of Finance, 54(4), 1361–1395.
- Ait-Sahalia, Y. (2002). Maximum likelihood estimation of discretely sampled diffusions: A closed-form approximation approach. Econometrica, 70(1).

Chan, K., Karolyi, G., Longstaff, F., & Sanders, A. (1992). An empirical comparison of alternative models for the short-term interest rate. Journal of Finance, 47, 1209–1228.