A Quick Look at Short Rate Models
Cox-Ingersoll-Ross Model
and
Jump Diffusion CIR Model

Jesse Freitag

1 Introduction

There are many models in academic and industry literature that discuss ways to model the evolution of interest rates. The short rate refers to the instantaneous interest rate on a very short-term risk-free borrowing or lending. It is a construct used to model the evolution of interest rates over time and forms the basis for many term structure models. The short rate r_t is the interest rate applicable over an infinitesimally small time interval, $[t, t+\delta]$. Unlike longer-term rates (e.g., 1-year, 10-year rates), which are averages or fixed over a term, the short rate evolves continuously in models.

The short rate is central to term structure models because all bond prices (or yields) can, in theory, be derived from the short rate. The price of a zero-coupon bond is often expressed as:

$$P(t,T) = \mathbb{E}_t\left[\exp\left(-\int_t^T r_s ds\right)\right]$$

where $P(t,T)$ is the price of a bond at time t maturing at T [5].

Among the more popular short rate models comes the Cox-Ingersoll-Ross (CIR) model [3]. Introduced in 1985 by John C. Cox, Jonathan Ingersoll, and Stephen Ross, it is an extension of the Vasicek model – itself an Ornstein-Uhlenbeck process, or mean-reversion model.

The model is given by the following stochastic differential equation [3]:

$$dr_t = a(b - r_t)dt + \sigma\sqrt{r_t}d\mathbf{W}_t$$

where:

- r_t is the instantaneous interest rate at time t
- $a > 0$ is the speed of mean reversion
- $b > 0$ is the long-term mean level of the rate
- $\sigma > 0$ is the volatility
- $\mathbf{W}_t$ is a Wiener process (Brownian motion)

2 Properties of the CIR Model

2.1 Mean Reversion

When $r_t > b$ i.e., the rate is above the mean, the term $b - r_t$ becomes negative. The drift term $a(b - r_t)$ pulls the rate downward, reducing r_t over time.
Conversely, when $r_t < b$, i.e., the rate is below the mean, the term $b - r_t$ becomes positive. The drift term $a(b - r_t)$ pulls the rate upward, increasing r_t over time. [4]
In the third case, when $r_t = b$, the drift $a(b - r_t) = 0$ and there is no pull toward or away from b. The rate may still fluctuate due to the stochastic term $\sigma\sqrt{r_t}dW_t$, but it does not systematically drift away from b.

Note that because of the term $\sigma\sqrt{r_t}$ in the stochastic component of the SDE, the CIR model never allows for negative interest rates and is bounded by 0. Short rate models like the Vasicek model do not have a geometric component in the stochastic component, thus allowing for rates to become negative. An interest rate of zero is also precluded if the condition $2ab \geq \sigma^2$ (Feller Condition) is met. More generally, when the rate r_t is close to zero, the standard deviation $\sigma\sqrt{r_t}$ also becomes very small which dampens the effect of the random shock on the rate.

To analyze the boundary behavior near $r_t = 0$, we consider the generator applied to a test function $f(r)$. The infinitesimal generator acting on $f(r)$ is:

$$\mathcal{L}f(r) = a(b-r)f'(r) + \frac{1}{2}\sigma^2 r f''(r)$$

When $r_t \to 0$, the drift term is $a(b-r)f'(r)$. As r_t approaches zero, this term becomes $abf'(r)$, which is a constant. As aforementioned, this term pushes r_t upwards and is not dampened by the value of r_t, making it a stabilizing force near zero.
Turning our heads to the diffusion term,$\frac{1}{2}\sigma^2 r_t f''(r)$, as r_t approaches zero, this term vanishes since $r_t \to 0$. The diffusion term represents the random shocks, but because $\sigma\sqrt{r_t} \to 0$ as $r_t \to 0$, the influence of randomness diminishes near zero. Hence, the effect of the stochastic component is reduced at low interest rates.

The key to understanding the boundary behavior of the CIR process at $r_t = 0$ lies in the Feller condition [2]. The Feller condition states that for the process to remain positive and never hit zero, the following condition must hold:

$$2ab \geq \sigma^2$$

If this condition is satisfied, the boundary at $r_t = 0$ is non-attracting, meaning the process will not tend toward zero as time progresses. The drift term $a(b-r)$ dominates the diffusion term near $r_t = 0$, pushing the process upward, ensuring that the short rate stays positive and is bounded away from zero.

If the condition is violated (i.e., if $2ab < \sigma^2$), the boundary at $r_t = 0$ becomes attracting, meaning the process can potentially reach zero in finite time. In this case, r_t could hit zero and remain stuck at zero, which can cause numerical instability in simulations.

2.2 Speed of Mean Reversion (a)

The Mean-reversion speed (measured in inverse time, e.g., per year) is denoted by a. A high a means the rate quickly reverts to b, leading to less variablility around the mean. Conversely, low a reverts more slowly, allowing for prolonged deviations from b.

A useful analogy of mean reversion for this model, as well as other Ornstein-Uhlenbeck processes is imagining the short rate as a ball rolling on a curve shaped like a bowl, where b is the bottom of the bowl (equilibrium) and a determining how steep or flat the bowl is. A steep bowl (high a) pulls the ball back to b quickly. A flat bowl (low a) allows the ball to linger away from b. The stochastic term $\sigma\sqrt{r_t}dW_t$ adds random "kicks" to the ball, causing it to oscillate around b instead of settling at it exactly.

2.3 Finding an explicit solution for r_t

I will show a derivation to find an explicit solution for r_t using Ito's Lemma.

Define the integrating factor:

$$X_t = e^{at} r_t$$

Applying Ito's Lemma to X_t, we differentiate:

$$dX_t = e^{at}dr_t + ae^{at}r_tdt$$

Substituting dr_t from the CIR SDE:

$$dX_t = e^{at}[a(b - r_t)dt + \sigma\sqrt{r_t}dW_t] + ae^{at}r_tdt$$

$$\implies dX_t = abe^{at}dt - ae^{at}r_tdt + \sigma e^{at}\sqrt{r_t}dW_t + ae^{at}r_tdt$$

Since the $-ae^{at}r_tdt$ and $ae^{at}r_tdt$ cancel, we are left with:

$$dX_t = abe^{at} + \sigma e^{at}\sqrt{r_t}dW_t$$

Integrating both sides from 0 to t:

$$X_t = X_0 + \int_0^t abe^{as}ds + \int_0^t \sigma e^{as}\sqrt{r_s}dW_s$$

Since $X_0 = r_0$ the first integral evaluates as :

$$\int_0^t abe^{as}ds = ab\int_0^t e^{as}ds = ab\frac{e^{at} - 1}{a} = b(e^{at} - 1)$$

Thus,

$$X_t = e^{at}r_0 + b(e^{at} - 1) + \sigma\int_0^t e^{as}\sqrt{r_s}dW_s$$

Finally, rewriting in terms of r_t:

$$r_t = e^{-at}r_0 + b(1 - e^{-at}) + \sigma e^{-at}\int_0^t e^{as}\sqrt{r_s}dW_s$$

Taking expectations,

$$\mathbb{E}[r_t] = e^{-at}r_0 + b(1 - e^{at}) + \sigma e^{-at}\mathbb{E}\left[\int_0^t e^{as}\sqrt{r_s}dW_s\right]$$

Since the Ito integral has zero mean, the above simplifies to:

$$\mathbb{E}[r_t] = e^{-at}r_0 + b(1 - e^{-at})$$

This confirms that r_t mean-reverts to b over time. [3]

3 Simulating CIR Model

The continuous SDE can be discretized as follows (useful for simulating in Python):

$$r_{t+\Delta t} - r_t = a(b - r_t)\Delta + \sigma\sqrt{r_t\Delta t}\epsilon_t$$

provided that $\epsilon_t \sim N(0, 1)$
Algorithm 3.1 exhibits pseudocode for simulating interest rate paths using the Cox-Ingersoll-Ross (CIR) model.

Figure 1 shows 50 simulated paths of interest rates (notice the strictly positive values?) using the parameters in Table 3.1:

Algorithm 3.1 Simulate Cox-Ingersoll-Ross (CIR) Interest Rate Paths

Require: a (mean reversion speed), b (long-term mean), σ (volatility), r_0 (initial rate), T (time horizon), Δt (time step), m (number of paths)

$n \leftarrow T/\Delta t + 1$
Set initial rates r_0
for $t \leftarrow 1$ to $n-1$ **do**
    Generate $z \leftarrow$ standard normal random variables of size m
    Compute drift $\leftarrow a \cdot (b - r_{t-1}) \cdot \Delta t$
    Compute diffusion $\leftarrow \sigma \cdot \sqrt{r} \cdot \sqrt{\Delta t} \cdot z$
    Update $r_t \leftarrow$r$_{t-1}$+ drift + diffusion
    Return *rates*

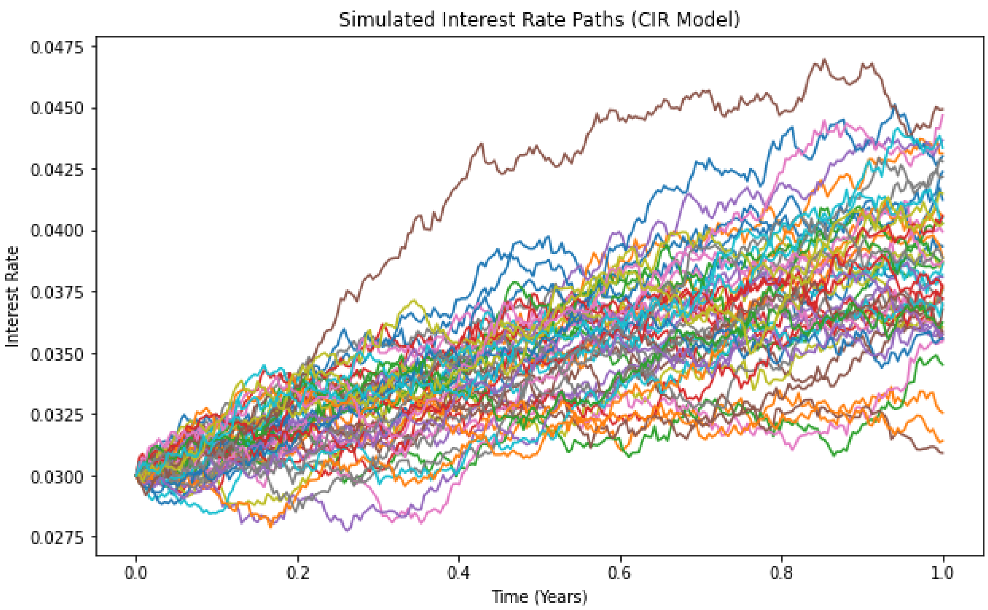

Figure 1 50 Simulated Paths from CIR Model

Table 3.1 Parameters for CIR Model Simulation

Parameter	**Value**	**Description**
a	0.5	Speed of mean reversion
b	0.05	Long-term mean interest rate
σ	0.02	Volatility parameter
r_0	0.03	Initial interest rate
T	1.0	Time horizon (1 year)
Δt	$\frac{1}{252}$	Time step (daily; 252 trading days/year)
m	50	Number of simulated paths

4 Jump Diffusion Cox Ingersoll Ross (CIR) Model

The Cox-Ingersoll-Ross model uses a special case of a basic affine jump diffusion – permitting a closed-form expression for bond prices. A basic affine jump diffusion process is a stochastic process,

Z takes the form of [1]

$$dZ_t = \kappa(\theta - Z_t)dt + \sigma\sqrt{Z_t}dB_t + dJ_t, \quad t \geq 0, Z_0 \geq 0$$

where B is the standard Brownian motion and J is an compound Poisson jump process with intensity λ, meaning jumps occur at a rate of λ per unit time. When a jump occurs, the jump size, $\mathbf{Y}$ is drawn from an exponential distribution following $\mathbf{Y} \backsim \mathrm{Exp}(\mu^{-1})$ The above looks familiar, right?

To understand how jumps affect the expected path of interest rates, we derive $\mathbb{E}[r_t]$, the expectation of the short rate under the jump-diffusion extension [1].

$$d\mathbb{E}[r_t] = \mathbb{E}[a(b - r_t)]dt + \mathbb{E}[\sigma\sqrt{r_t}dW_t] + \mathbb{E}[dJ_t]$$

Because dW_t has zero mean and it independent of r_t, its expectation vanishes. We are left with:

$$d\mathbb{E}[r_t] = a(b - \mathbb{E}[r_t])dt + \mathbb{E}[dJ_t]$$

For the jump term, the expectation is given by:

$$\mathbb{E}[dJ_t] = \lambda\mathbb{E}[Y]dt = \lambda\mu dt$$

$$\implies d\mathbb{E}[r_t] = a(b - \mathbb{E}r_t)dt + \lambda\mu dt$$

The above is a first-order linear ODE, which we can solve by defining $R_t = \mathbb{E}[r_t]$

$$\frac{dR_t}{dt} + aR_t = ab + \lambda\mu$$

Using the integrating factor e^{at}, we multiply both sides:

$$e^{at}\frac{dR_t}{dt} + ae^{at}R_t = e^{at}(ab + \lambda\mu)$$

$$\implies \frac{d}{dt}(e^{at}R_t) = e^{at}(ab + \lambda\mu)$$

Then integrating from 0 to t :

$$\int_0^t \frac{d}{dt}(e^{at}R_t) = \int_0^t e^{as}(ab + \lambda\mu)ds$$

$$\implies e^{at}R_t - R_0 = (ab + \lambda\mu)\int_0^t e^{as}ds$$

$$e^{at}R_t - R_0 = (ab + \lambda\mu)\frac{e^{at} - 1}{a}$$

[1]

$$R_t = e^{at}R_0 + \left(b + \frac{\lambda\mu}{a}\right)(1 - e^{at})$$

Thus the expected short rate under the jump diffusion CIR model is:

$$\mathbb{E}[r_t] = r_0e^{-at} + \left(b + \frac{\lambda\mu}{a}\right)(1 - e^{at})$$

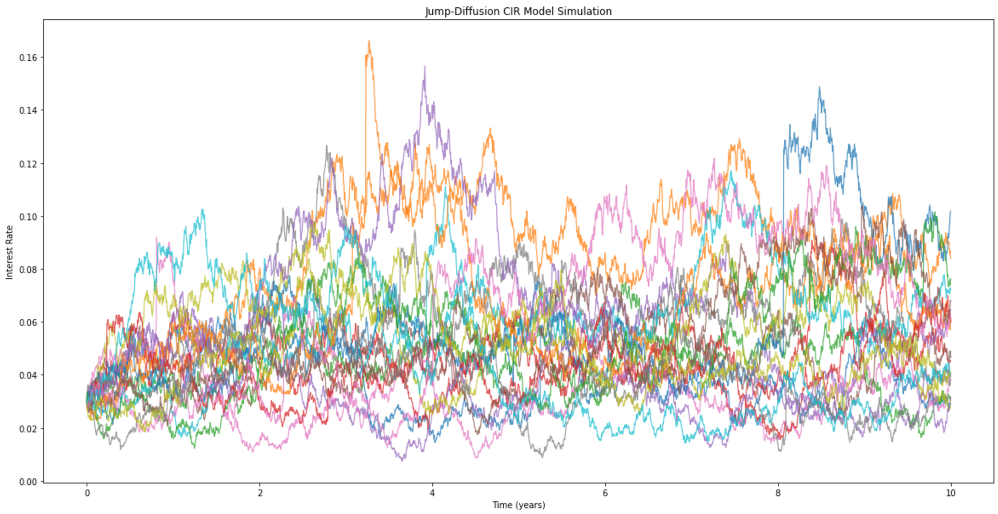

Figure 2 Jump Diffusion CIR Model

This result shows that the mean short rate exhibits mean reversion, but now toward an adjusted level:

$$b^* = b + \frac{\lambda\mu}{a}$$

where the addition of the fractional term means that persistent positive jumps increase the long-term equilibrium interest rate.

Simulating a Jump Diffusion CIR process in Python is a lot similar to the initial CIR model-simulation in Figure 1. See **Figure 2** for the Jump Diffusion simulation using similar parameters, adding a $\mu = 0.02$ for the mean jump size and $\lambda = 0.2$ for the Poisson jump intensity.

5 Parameter Estimation and Calibration for the Jump Diffusion Model

Estimating the parameters of a Jump-Diffusion CIR model is crucial for pricing any interest rate derivative. It is not anywhere as useful simply 'guessing' any parameters and hitting compile in your IDE despite the nicely looking plots.

A very common way to estimate the parameters $\theta = (a, b, \sigma, \lambda, \mu)$ of such model is using historical short rate data and apply one of the following approaches – Maximum Likelihood Estimation or Generalized Method of Moments. Some approaches estimate the parameters using both methods and then taking the average of the estimates for their model. I will now show how to estimate using both methods.

5.1 Maximum Likelihood Estimation

The log-likelihood function for the CIR process without jumps:

$$f(r_t|r_{t-1}) = \frac{c}{p^2}\exp\left(-\frac{c(r_t e^{-a\Delta t} - b(1 - e^{-a\Delta t}))^2}{2\sigma^2(1 - e^{-2a\Delta t})}\right)$$

where $c = \frac{2a}{\sigma^2(1-e^{-2a\Delta t})}$ For the jump component, we modify the likelihood using a mixture model, integrating the Poisson process's probability mass function.

For the **Jump Component Likelihood** the number of jumps $N_t \backsim$ **Poisson**$(\lambda \Delta t)$ and jumps are i.i.d exponentially distributed with mean μ. Denote the following:

- Probability of no jump : $P(N_t = 0) = e^{-\lambda \Delta t}$
- Probability of one jump : $P(N_t = 1) = (\lambda \Delta t) e^{-\lambda \Delta t}$
- The jump size follows $J \sim \text{Exp}(1/\mu)$

Thus, the mixture likelihood function (combining CIR + Jumps) is:

$$\mathcal{L}(r_t | r_{t-1}) = e^{-\lambda \Delta t} f_{\text{CIR}}(r_t | r_{t-1}) + (1 - e^{-\lambda \Delta t}) f_{\text{Jump}}(r_t)$$

where

$$f_{\text{Jump}}(r_t) = \frac{1}{\mu} e^{-r_t/\mu}$$

Taking the log-likelihood over all observations:

$$\log \mathcal{L} = \sum_{t=1}^{T} \log\Big(e^{-\lambda \Delta t} f_{\text{CIR}}(r_t | r_{t-1}) + (1 - e^{-\lambda \Delta t}) f_{\text{Jump}}(r_t)\Big)$$

5.2 Generalized Method of Moments (GMM)

Continuing on with the second parameter calibration method, GMM is a method that estimates parameters by matching empirical and theoretical moments. For the Jump Diffusion CIR model, we have the SDE that we have seen a few times now:

$$dr_t = a(b - r_t)dt + \sigma \sqrt{r_t} dW_t + dJ_t$$

To apply GMM, we derive theoretical moments of r_t incorporating the mean, variance, and higher-order moments [1]. We have already shown that the mean is

$$\mathbb{E}[dr_t] = a(b - \mathbb{E}[r_t])dt + \mathbb{E}[dJ_t] \implies \mathbb{E}[r] = \frac{ab + \lambda \mu}{a}$$

To calculate the variance, we use Ito's Lemma:

$$d(r_t^2) = 2 r_t dr_t + (dr_t)^2$$

Using $(dr_t)^2 = \sigma^2 r_t dt$ the variance equation is:

$$\frac{d\text{Var}(r_t)}{dt} = 2a(b - \mathbb{E}[r_t])\mathbb{E}[r_t] + \sigma^2 \mathbb{E}[r_t] + \lambda\Big(\mu^2 + 2\mu \mathbb{E}[r_t]\Big)$$

At equilibrium:

$$\text{Var}(r_t) = \frac{\sigma^2 \mathbb{E}[r_t] + \lambda(\mu^2 + 2\mu \mathbb{E}[r_t])}{2a}$$

For the third moment, skewness, its a bit long to include the derivation here, but we have:

$$\text{Skewness}(r_t) = \frac{\mathbb{E}[(r_t - \mathbb{E}[r_t]^3)]}{\text{Var}(r_t)^{3/2}}$$

which captures jump-induced asymmetry. From here, we calculate the empirical moments from historical data:

Empirical Mean:

$$\widehat{\mathbb{E}}[r] = \frac{1}{T}\sum_{t=1}^{T} r_t$$

Empirical Variance:

$$\widehat{\mathrm{Var}}(r) = \frac{1}{T}\sum_{t=1}^{T}(r_t - \widehat{\mathbb{E}}[r])^2$$

Empirical Skewness:

$$\widehat{\mathrm{Skew}}[r] = \frac{1}{T}\sum_{t=1}^{T}\left(\frac{r_t - \widehat{\mathbb{E}[r]}}{\sqrt{\widehat{\mathrm{Var}(r)}}}\right)^3$$

Then we define the moment conditions:

$$g(\theta) = \begin{bmatrix} \mathbb{E}[r] - \widehat{\mathbb{E}}[r] \\ \mathrm{Var}(r) - \widehat{\mathrm{Var}(r)} \\ \mathrm{Skewness}(r) - \widehat{\mathrm{Skewness}}(r) \end{bmatrix}$$

The GMM objective function minimizes:

$$J(\theta) = g(\theta)' W g(\theta)$$

where W is a weighting matrix. Popularly, setting $W = \mathbf{I}$, the identity matrix, then GMM assumes equal weighting for all moment conditions.

6 Conclusion

Although quite difficult, speaking from personal experience, due to model development choices, the parameters of the models can be calibrated either by Maximum Likelihood Estimation or the Generalized Method of Moments, or an average of the two.

To conclude, this was a short article showing the dynamics and behavior of the Cox-Ingersoll-Ross (CIR) model and its extensions, including the jump diffusion model, which are very useful for modeling the evolution of short-term interest rates. The CIR model's mean-reverting nature and its ability to prevent negative interest rates make it a reliable choice for interest rate modeling in finance over other simpler models like the Vasicek model. By incorporating jump diffusion, we account for the possibility of sudden shocks in interest rates.

References

[1] Federal Reserve Board. Jump-diffusion term structure and ito conditional moment generator. Technical Report 2001-28, Federal Reserve Board, 2001. URL https://www.federalreserve.gov/pubs/feds/2001/200128/200128pap.pdf.

[2] M. Jeanblanc. *Mathematical Methods for Financial Markets*. Springer Finance, 2009.

[3] S. A. R. John C Cox, Jonathan E Ingersoll Jr. A theory of the term structure of interest rates. March 1985. URL https://www.econometricsociety.org/publications/econometrica/1985/03/01/theory-term-structure-interest-rates.

[4] S. E. Shreve. The dynamics of the cox-ingersoll-ross process. *Handvbook of Mathematical Finance*, 2003.

[5] B. Tuckman. *Fixed Income Securities: Tools for Today's Markets*. John Wiley & Sons, 2002.