What Short Rate Model Should I Use？

Colin Turfus
Aurelio Romero-Bermudez

Wilmott Magazine

Aki Lin

Mar 31, 2022

Abstract

- Choosing the best short rate model combining analytic tractability with accurate representation of the term structure of interest rate volatility
- suggest a new approach which provides approximate but highly accurate analytic bond and option pricing while allowing a trade-off between complexity and goodness-of-fit to the market-observed term structure of smile and skew

Outline

- Introduction
- Overview of Short Rate Models
- Canonical Framework
- Functional Forms for the Short Rate
- Stochastic Discounting Operators
- Numerical Evidence
- Conclusions

Introduction

- What is the best short rate model to represent interest rates?
- What the criteria are to be used in determining ”best”?
- Consider a wider range is likely to give rise to a better answer.

Overview of Short Rate Models

- Classical Models
  - generic one-factor dynamics for the short rate

$$dr_t = \mu(r_t, t)dt + \sigma(r_t, t)dW_t$$

  - the volatility $\sigma(r_t, t)$ is an affine function of r_t
  - $\sigma(r_t, t)$
    - [1985] Cox et al. $\mathrm{d}r(t) = k(\theta - r(t))\mathrm{d}t + \sigma\sqrt{r(t)}\mathrm{d}W(t)$
      - the volatility $\sigma^2(r_t, t)$ is an affine function of r_t
  - $\mu(r_t, t)$
    - [1978] Dothan $\mathrm{d}r(t) = ar(t)\mathrm{d}t + \sigma r(t)\mathrm{d}W(t)$
    - [1990] Hull and White $dr(t) = [\theta(t) - \alpha(t)r(t)]\ dt + \sigma(t)\,dW(t)$
    - [1991] Black and Karasinski $d\ln(r) = [\theta_t - \phi_t \ln(r)]\,dt + \sigma_t\,dW_t$
    - [2001] Mercurio and Moraleda $dr(t) = [\theta(t) - \beta(t)r(t)]\,dt + \sigma(t)dW(t)$

Overview of Short Rate Models

- Classical Models
  - generic one-factor dynamics for the short rate

    $$dr_t = \mu(r_t, t)dt + \sigma(r_t, t)dW_t$$

  - the volatility $\sigma(r_t, t)$ is an affine function of r_t
  - $\sigma(r_t, t)$
    - [1985] Cox et al. $\mathrm{d}r(t) = k(\theta - r(t))\mathrm{d}t + \sigma\sqrt{r(t)}\mathrm{d}W(t)$
      - the volatility $\sigma^2(r_t, t)$ is an affine function of r_t
  - $\mu(r_t, t)$
    - [1978] Dothan $\mathrm{d}r(t) = ar(t)\mathrm{d}t + \sigma r(t)\mathrm{d}W(t)$
    - [1990] Hull and White $dr(t) = [\theta(t) - \alpha(t)r(t)]\ dt + \sigma(t)\,dW(t)$
    - [1991] Black and Karasinski $d\ln(r) = [\theta_t - \phi_t \ln(r)]\,dt + \sigma_t\,dW_t$
    - [2001] Mercurio and Moraleda $dr(t) = [\theta(t) - \beta(t)r(t)]\,dt + \sigma(t)dW(t)$

Overview of Short Rate Models

- Hull-White approach
  - While it was suggested that, in a low rates regime with negative rates possible or even common, there were certain advantages, these had to be weighed against the disadvantage that, in the event that the rates simulated by the model rose significantly, the modelled volatility did not, resulting in poor modelling for OTM call options.
- CIR
  - a hard lower bound for the simulated short rate with the risk of rates getting stuck there for extended periods with the volatility underestimated
  - the integral of the short rate to be incorporated is not known to be amenable.
- Black-Karasinski model
  - provides a more plausible alternative with volatility automatically rising when the simulated rates move away from zero.
  - the requirements of a hard minimum value for simulated rates and of the implied normal volatility vanishing as rates approach this minimum.

Overview of Short Rate Models

- The Black Shadow Rate Model
  - allowed rates to become negative
  - specifying that any rates which were modelled as negative were set to zero
- Local Volatility Models
  - have a more general dependence on rt than the relatively simple (typically affine) assumptions made in the classical models
- Stochastic Volatility Models
  - a stochastic volatility extension to the Hull-White model
  - $$dr_t = \alpha_r(t)(\theta_r(t) - r_t)dt + \sqrt{v_t}dW_t^1$$
    $$dv_t = \alpha_v(t)(\theta_v(t) - v_t)dt + \sqrt{v_t}dW_t^2$$
    $$\langle dW_t^1, dW_t^2 \rangle = \rho dt.$$
- Multi-Factor Models
  - cost of added complexity
  - facilitates greater flexibility in the taxonomy of shapes which a curve can take on in subsequent evolution, with different parts of the curve moving in different directions at the same time, rather than all forward rates being perfectly correlated as they are in the one-factor case

Canonical Framework

- In order to compare and evaluate models, it is useful to introduce a canonical framework which allows the essential features of a model to be seen more transparently.
  - the dependence of volatility on the short rate.
  - the dependence of mean reversion on the short rate.
- There is clearly never going to be one “best” model: our interest is in making a choice which is optimal against a range of criteria, analytic tractability being salient among them.

Canonical Framework

- In order to compare and evaluate models, it is useful to introduce a canonical framework which allows the essential features of a model to be seen more transparently.
  - the dependence of volatility on the short rate.
  - the dependence of mean reversion on the short rate.
- To facilitate this exploration
  - follow Martin et al. [2015] in making use of a Lamperti transform to remove the functional dependence of the volatility on the underlying rate.

$$dy = \frac{dr}{\sigma(r,t)}$$

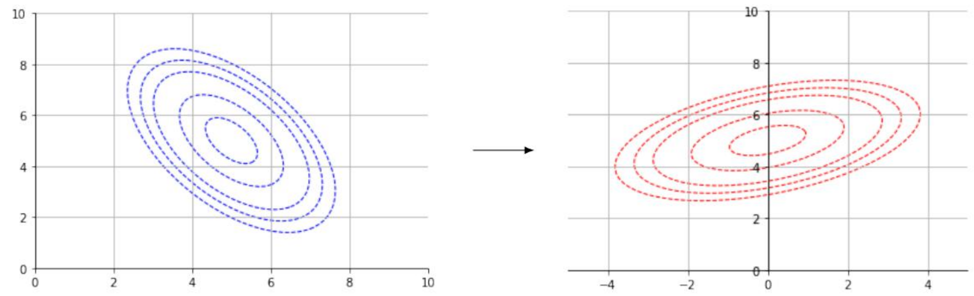

Figure 2.1: Exemplary figure to display the mapping of a process' state-space to a new space, that occurs when applying the Lamperti transform.

Canonical Framework

- Applying this process to the classical models, we see the following canonical structure arises:

$$
\begin{aligned}
dy_t &= g(y)(\theta(t) - \alpha(t)y_t)dt + \sigma(t)dW_t, \\
r_t &= r(y_t, t),
\end{aligned}
$$

- CIR :

$$g(y) = 1/y \text{ (with non-zero } \theta(t))$$

Functional Forms for the Short Rate

- We observe that there appears to have been surprisingly little variety in the choices which have been considered (implicitly or explicitly) in the literature for the function $r(y,t)$

$$r(y,t) := a(t) + b(t)y + c(t)y^2.$$

- it has analytic bond prices, but the time-dependent coefficients arising in the bond formula are obtained by solving non-linear Ricatti ODEs.
- Option prices are then obtained by numerical integration of the payoff function over the Gaussian variable y
- not a monotonic function of y, so does not give rise to a bijection between rt and yt

Functional Forms for the Short Rate

- The property that rt is bounded above by U and below by L
- They provide a means of deriving approximate asymptotic expressions for zero coupon bond prices, valid in the limit of low volatility

$$r(y,t) := \frac{U(t)e^{\beta(t)y} + L(t)\alpha(t)}{e^{\beta(t)y} + \alpha(t)}$$

- lognormal volatilities tend to be high, of the order of 60%, so a low volatility assumption is unlikely to sit well with this important use case.

Functional Forms for the Short Rate

- This model is of particular interest since, for different choices of the β parameter, it gives rise to a family which contains both the Black-Karasinski and the Hull-White families of models as special cases: the former arising with $\beta := 0$ and the latter arising in the limit as β approach to 1.
- With different values of β giving rise to different skewness levels

$$r(y,t) := R(t)\frac{\exp\frac{(1-\beta(t))y}{R(t)^{\beta(t)}} - \beta(t)}{1-\beta(t)}.$$

- Horvath et al. [2017] derive highly accurate asymptotic expressions for bond and option prices, subject only to interest rate fluctuations being small in absolute terms; which they invariably are, even in the highest interest rate régimes.

Functional Forms for the Short Rate

- if one extends the Hull-White model by introducing a quadratic local volatility distribution

$$\sigma(x_t, t) = \sigma(t)\sqrt{1 + \gamma^2(t)(x_t + y^*(t))^2}$$

- applies a Lamperti transform

$$r(y,t) := \overline{r}(t) + R^*(t) + \frac{\sinh \gamma(t)(y + y^*(t))}{\gamma(t)}.$$

- Having now three configurable parameters $\sigma(t)$ $y^*(t)$ $\gamma(t)$
- Although the resultant model does not have exact analytic bond and option solutions, it is possible to derive using the methods expounded by Turfus [2021a] highly accurate asymptotic representations, valid subject only to the constraint that again rate fluctuations are not high.

Functional Forms for the Short Rate

- Turfus and Romero-Bermúdez [2021]

$$r(y,t) := \bar{r}(t) + C_0(t) + \sum_{n=1}^{N} C_n(t) e^{\gamma_n(t) y}$$

$$\textstyle\sum_{n=1}^{N} C_n(t) = 0$$

$$\textstyle\sum_{n=1}^{N} C_n(t)\gamma_n(t) = 1.$$

- $C_n(t)$ and $\gamma_n(t)$ equal but with opposite signs.

- This would give rise to an additional degree of freedom, facilitating different limiting smile levels for low and high rates, with almost no concomitant increase in model complexity

Stochastic Discounting Operators

- We therefore look now at the way in which solutions for the canonical short rate model specified by (3.3) and (4.6) can be obtained systematically
- To this end we introduce the idea of stochastic discounting operators and use them to obtain pricing formulae for one model as perturbations of those for another.

$$r_t = r(y_t, t), \tag{3.3}$$

$$dy_t = -\alpha(t) y dt + \sigma(t) dW_t. \tag{4.6}$$

- backward Kolmogorov equation

$$\left(\frac{\partial}{\partial t} + \mathcal{L}(t)\right) f(y,t) = 0,$$

$$\mathcal{L}(t) := -\alpha(t) y \frac{\partial}{\partial y} + \frac{1}{2}\sigma^2(t)\frac{\partial^2}{\partial y^2} - r(y,t)$$

- In seeking approximate analytic solutions, the trick is, in the first instance, to follow Hagan et al. [2015] in writing

$$\mathcal{L}(t) = \mathcal{L}_0(t) + \mathcal{V}(t)$$

Stochastic Discounting Operators

- $\mathcal{L}_0(t)$ gives rise to a well-known integral kernel with a convenient analytic representation
- $\mathcal{V}(t)$ is small in the sense that its evolution operator (as defined below) can be expressed as a rapidly convergent series
- The integral kernel or Green's function has the property that

$$f(y,t) = \int_{\mathbb{R}} G(y,t;\eta,T)P(\eta)d\eta.$$

$$G(y,t;\eta,T) = \mathcal{E}_t^T(-\mathcal{W}(t,\cdot))G_0(y,t;\eta,T).$$

$$G_0(y,t;\eta,T) = D(t,T)N\left(\frac{\eta - \phi_y(t,T)y}{\sqrt{\Sigma_{yy}(t,T)}}\right)$$

$$\mathcal{D}(t,T) = \mathcal{E}_t^T(-\mathcal{W}(t,\cdot))$$

Stochastic Discounting Operators

- $\mathcal{L}_0(t)$ gives rise to a well-known integral kernel with a convenient analytic representation
- $\mathcal{V}(t)$ is small in the sense that its evolution operator (as defined below) can be expressed as a rapidly convergent series
- Notation

$$D(t,v) := e^{-\int_t^v \overline{r}(u)\,du}, \tag{5.10}$$

$$\phi_y(t,v) := e^{-\int_t^v \alpha(u)du}, \tag{5.11}$$

$$B^*(t,v) := \int_t^v \phi_y(t,u)du, \tag{5.12}$$

$$z_t = \int_0^t (r(y_u,u) - \overline{r}(u))du \tag{5.13}$$

$$\Sigma_{yy}(t,v) := \int_t^v \phi_y^2(u,v)\sigma^2(u)du, \tag{5.14}$$

$$\Sigma_{yz}(t,v) := \int_t^v \phi_y(u,v)\Sigma_{yy}(t,u)du, \tag{5.15}$$

$$\Sigma_{zz}(t,v) := 2\int_t^v \Sigma_{yz}(t,u)du, \tag{5.16}$$

$$\mu^*(y,t,v) := B^*(t,v)(y + \Sigma_{yz}(0,t)) + \tfrac{1}{2}B^{*2}(t,v)\Sigma_{yy}(0,t). \tag{5.17}$$

Stochastic Discounting Operators

- Hull-White
  - For the Hull-White model we have $r(y,t) = \overline{r}(t) + y + \Sigma_{yz}(0,t)$

$$\mathcal{W}(t,u) = \phi_y(t,u)y + \Sigma_{yz}(0,t) + \frac{\Sigma_{yy}(t,u)}{\phi_y(t,u)}\frac{\partial}{\partial y}.$$

$$\mathcal{D}(t,T) = e^{-\mu^*(y,t,T)}\mathcal{M}^-(t,T)$$

$$\mathcal{M}^{\pm}(t,u)(f(y,\ldots)) = f(y \pm \Delta y(t,u),\ldots),$$

$$\Delta y(t,u) := \frac{\Sigma_{yz}(t,u)}{\phi_y(t,u)}.$$

  - In this case the time-ordered exponential can be calculated and applied

$$G_{\text{H-W}}(y,t;\eta,T) = e^{-\mu^*(y,t,T)}G_0(y - \Delta y(t,T), t; \eta, T).$$

Stochastic Discounting Operators

- Black-Karasinski
  - we have

$$r(y,t) = \tilde{r}(t)e^{y-\frac{1}{2}\Sigma_{yy}(0,t)}$$

$$\mathcal{W}(t,u) = \tilde{r}(t)e^{\phi_y(t,u)y-\frac{1}{2}\phi_y^2(t,u)\Sigma_{yy}(0,t)}\,\mathcal{M}^+(t,u) - \overline{r}(t).$$

$$\Delta y(t,u) := \frac{\Sigma_{yy}(t,u)}{\phi_y(t,u)}$$

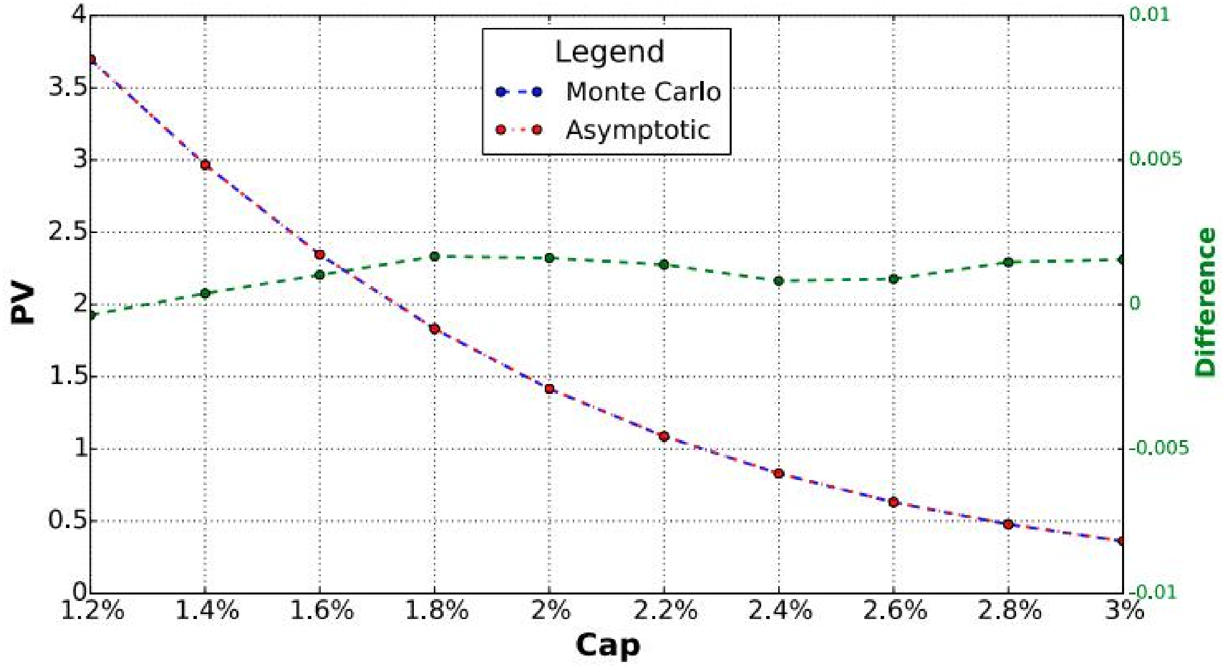

Figure 2: Accuracy of analytic cap prices for Black-Karasinski model

Stochastic Discounting Operators

- Sinh Model
-

$$\mathcal{W}(t,u) = R^{+}(y,t,u)\,\mathcal{M}^{+}(t,u) - R^{-}(y,t,u)\,\mathcal{M}^{-}(t,u) + R^{*}(u),$$

$$R^{\pm}(y,t,u) := \frac{e^{\frac{1}{2}\gamma^{2}(u)\Sigma_{yy}(t,u)\pm\gamma(u)(\phi_{y}(t,u)y+y^{*}(u))}}{2\gamma(u)}$$

$$\Delta y(t,u) := \frac{\gamma(u)\Sigma_{yy}(t,u)}{\phi_{y}(t,u)}$$

Stochastic Discounting Operators

- Beta Blend
- Multi-Exponential Model
- Composite Kernel
- Compounded Rates Kernel
- Hybrid Kernels

Numerical Evidence

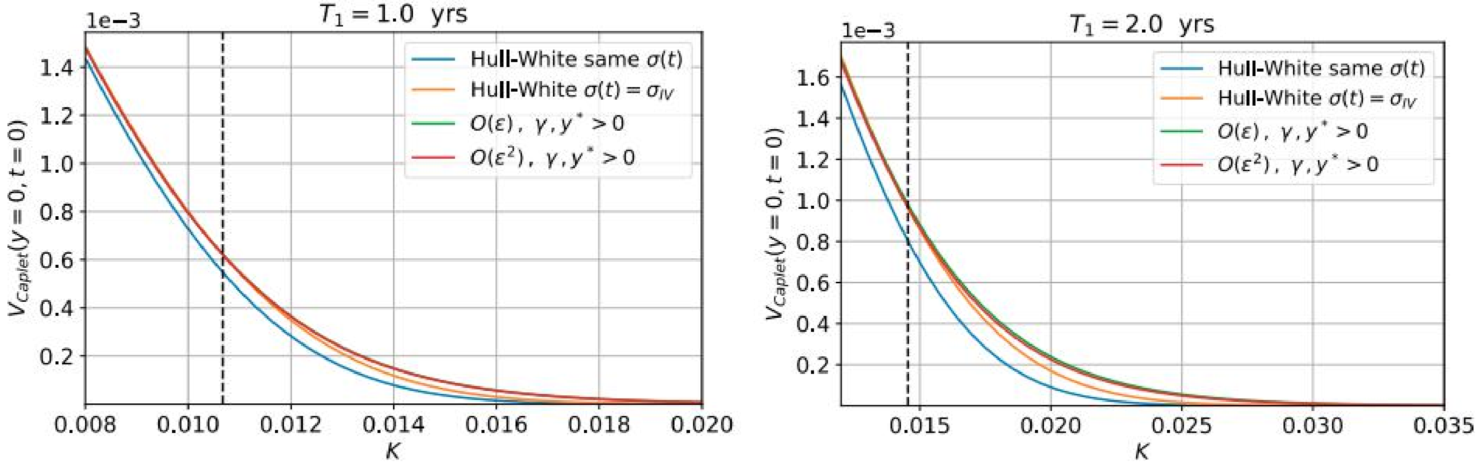

Figure 3: Caplet PV for various strikes K.

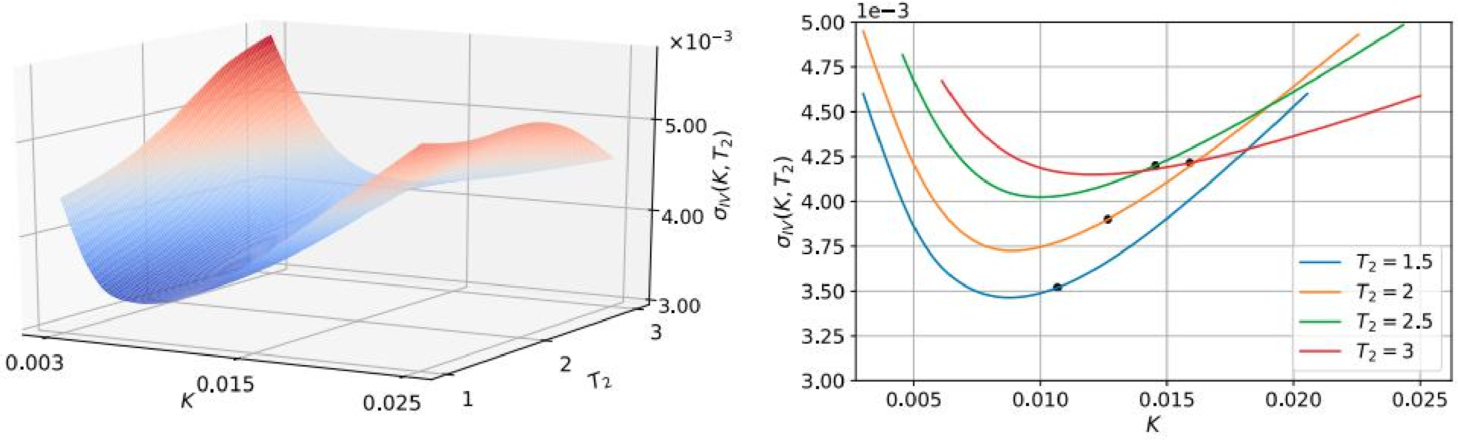

Figure 4: Implied Hull-White volatility surface

Conclusions

- We then turned our attention to some more recent models which give the ability potentially to fit the term structure of volatility smile and skew while preserving tractable analytic approximations (involving only quadratures) for bond and option prices.
- a more fruitful way forward to explore alternative short rate models is to represent the short rate as a function $r(y_t, t)$ of a mean-reverting Gaussian variable y_t

Thanks!

Aki Lin

Mar 31, 2022