*Quantitative Finance*, 2014
Vol. 14, No. 11, 1961–1970, http://dx.doi.org/10.1080/14697688.2013.765062

Routledge
Taylor & Francis Group

# An alternative approach to the calibration of the Vasicek and CIR interest rate models via generating functions

MARIANITO R. RODRIGO† and ROGEMAR S. MAMON*‡

†School of Mathematics and Applied Statistics, University of Wollongong, Wollongong, New South Wales, Australia
‡Department of Statistical and Actuarial Sciences, University of Western Ontario, London, Ontario, Canada

(*Received 20 October 2011; in final form 5 January 2013*)

We propose a new method to calibrate the Vasicek and Cox–Ingersoll–Ross interest rate models from bond prices. We define an appropriate generating function and derive recursive relations between the derivatives of the generating function and the bond prices. The parameters of the Vasicek and CIR models are then obtained by solving a system of linearly independent equations arising from the recursive relations. We include numerical results that show the method's accuracy when bond prices generated from the exact formulas are used.

*Keywords*: Interest rate models; Exponential affine bond price; Calibration method; Generating functions; Integral transforms

*JEL Classification*: C, G1

## 1. Introduction

The successful implementation of any financial model depends on being able to perform calibration, which is the process of determining the parameters by fitting the model to observed market data. This type of calibration, through the inversion of the yield curve, is driven by the principle of market-consistent valuation. Such a principle is supported by the fact that it is the market that chooses the martingale measure (Björk 2009). The problem is to solve for the set of parameters in the bond pricing formula that exactly gives the current market statistics.

This article presents an alternative calibration method of two popular short rate models, namely the Vasicek and Cox–Ingersoll–Ross (CIR) models. The fundamental object to calibrate against is the yield curve, which can be constructed from bond prices with different maturities. Calibration is important in the pricing of term-structure derivatives and other insurance products with option-linked characteristics, where risk-neutral parameters are required. Both the Vasicek and CIR models have been employed ubiquitously in the valuation of contracts dependent on the movement of the yield curve. These include, among others, the valuation of sinking-fund bonds (Bacinello *et al*. 1996), life insurance contracts (Biffis 2004), and contingent claims with correlated financial and mortality risks (Jalen and Mamon 2009).

As the interest rate process is characterized by a Gaussian distribution under the Vasicek model, it can become negative with a positive probability. This is considered as the significant drawback of the Vasicek model as a trade-off for its mathematical convenience. In fact, forward rates may also become negative for finite maturity. One may argue, however, that the probability of the interest rate becoming negative under the Vasicek framework can be controlled considering that the timescale is short or by having small volatility or putting a barrier near zero. Yet, under realistic circumstances and feasible calibration procedures, preventing the rates hitting negative values can only be done to a limited extent. Another major failing of the Vasicek model is that its specification does not align with the empirical observation that the short-term volatility is non-constant. The CIR model was then proposed as a tractable model while keeping the rates positive with volatility as an increasing function of the interest rate. But, as both the Vasicek and CIR models belong to the one-factor family of models, they are criticized for their inability to reproduce certain yield curve shapes that occur occasionally in practice. Despite some of the imperfections of the Vasicek and CIR models, they are Markovian models that can be implemented by lattice trees and employed for pricing path-dependent term-structure derivatives. Above all, they

*Corresponding author. Email: rmamon@stats.uwo.ca

© 2013 Taylor & Francis

are intuitive and the ease of interpretation of their dynamics as well as the availability of their closed-form bond price solutions merits their discussion and use for calibration, which is the ultimate goal of modeling implementation.

Before we develop our proposed calibration method and cite previous work on this problem, it would be worthwhile to briefly describe our framework. Such a framework is similar in principle to the one used in addressing the inverse problem of volatility estimation in equity option pricing, where option prices with different strikes and maturities are utilized to obtain the underlying variable's volatility. More information can be found in Björk (2009), for example. In our case, we tackle the inverse problem of estimating risk-neutral parameters using current bond prices with different maturities.

Denote the price of a zero-coupon bond at time $t$ with maturity $T$ by $p(t,T)$. Although currently there are a number of ways to model the term structure $\{p(t,T):0 \leq t \leq T, T \geq 0\}$, we follow the classical route by postulating that it depends on the short rate $r = \{r_t : 0 \leq t \leq T\}$, i.e. we assume that

$$p(t,T) = F(t, r_t; T) = F^T(t, r_t),$$

for some function $F^T = F^T(t,x)$ that verifies the condition $F^T(T,x) = 1$ for all $x$. The $r$-dynamics is specified by a stochastic differential equation (SDE)

$$\mathrm{d}r_t = \mu^P(t, r_t)\mathrm{d}t + \sigma^P(t, r_t)\mathrm{d}W_t^P,$$

where $\mu^P$ and $\sigma^P$ are deterministic functions of $t$ and $x$ and $W^P = \{W_t^P : 0 \leq t \leq T\}$ is a Wiener process with respect to the objective probability measure $P$. It can be shown that $F^T$ satisfies the partial differential equation

$$\frac{\partial F^T}{\partial t} + [\mu^P(t,x) - \lambda(t,x)\sigma^P(t,x)]\frac{\partial F^T}{\partial x} + \frac{1}{2}\sigma^P(t,x)^2\frac{\partial^2 F^T}{\partial x^2} - xF^T = 0,$$

where $\lambda = \lambda(t,x)$ is the so-called market price of risk. Moreover, it has a risk-neutral valuation representation

$$F^T(t,x) = E^Q_{t,x}\left[\mathrm{e}^{-\int_t^T r_s\,\mathrm{d}s}\right],$$

where the subscripts $t$ and $x$ denote that the expectation is to be taken given the $r$-dynamics

$$\mathrm{d}r_s = [\mu^P(s, r_s) - \lambda(s, r_s)\sigma^P(s, r_s)]\mathrm{d}s + \sigma^P(s, r_s)\mathrm{d}W_s^Q, \quad r_t = x,$$

and $W^Q = \{W_t^Q : 0 \leq t \leq T\}$ is a Wiener process with respect to the martingale measure $Q$. Let $\mu^Q = \mu^P - \lambda\sigma^P$ and $\sigma^Q = \sigma^P$, so that

$$\mathrm{d}r_s = \mu^Q(s, r_s)\mathrm{d}s + \sigma^Q(s, r_s)\mathrm{d}W_s^Q, \quad r_t = x.$$

Thus, the term structure, and the prices of all other derivatives with underlying $r$, is completely determined by specifying the $r$-dynamics represented by $\mu^Q$ and $\sigma^Q$ under the martingale measure $Q$. Henceforth, we shall assume that we are working under this martingale measure and drop the superscript $Q$. Note that we make no implication here that the parameters recovered from our approach would produce SDEs with strictly positive sample path values. This is because the approach is based on the modeling assumption (i.e. the Vasicek model). In fact, since interest rates can take negative values under the Vasicek assumption, there is even no guarantee of obtaining sensible estimated parameters for every market dataset. If such a situation arises, it is probably reasonable to infer that the dataset's generating process is not Vasicek. However, theoretically, if the dataset is the exact bond price data arising from the Vasicek SDE, then we are guaranteed exact parameters.

Affine term structure models are those where the bond price has the exponential affine form

$$p(t,T) = F^T(t, r_t) = \mathrm{e}^{A(t,T) - r_t B(t,T)}, \tag{1.1}$$

for some deterministic functions $A$ and $B$ of $t$ and $T$. It is known (Duffie and Kan 1996, Elliott and van der Hoek 2001) that if

$$\mu(t,x) = \alpha(t)x + \beta(t), \qquad \sigma(t,x)^2 = \gamma(t)x + \delta(t),$$

where $\alpha$, $\beta$, $\gamma$, and $\delta$ are deterministic functions of $t$, then the bond price has an exponential affine form. Two of the most popular short rate models that admit an affine term structure are due to Vasicek (1977),

$$\mathrm{d}r_t = a(b - r_t)\mathrm{d}t + \sigma\mathrm{d}W_t,$$

and Cox, Ingersoll and Ross (1985),

$$\mathrm{d}r_t = a(b - r_t)\mathrm{d}t + \sigma\sqrt{r_t}\mathrm{d}W_t.$$

Here $a$, $b$, and $\sigma$ are positive constants that represent the speed of mean reversion, the mean-reverting level, and the volatility, respectively. In the Vasicek model we have

$$\begin{aligned} A(t,T) &= A(T-t) \\ &= -\frac{\sigma^2}{4a}B(T-t)^2 + \left(b - \frac{\sigma^2}{2a^2}\right)[B(T-t) - (T-t)], \end{aligned} \tag{1.2a}$$

$$B(t,T) = B(T-t) = \frac{1}{a}[1 - \mathrm{e}^{-a(T-t)}], \tag{1.2b}$$

while for the CIR model we have

$$\begin{aligned} A(t,T) &= A(T-t) \\ &= \frac{2ab}{\sigma^2}\log\frac{2c\,\mathrm{e}^{(c+a)(T-t)/2}}{(c+a)\mathrm{e}^{c(T-t)} + (c-a)}, \end{aligned} \tag{1.3a}$$

$$B(t,T) = B(T-t) = \frac{2[\mathrm{e}^{c(T-t)}-1]}{(c+a)\mathrm{e}^{c(T-t)}+(c-a)}, \tag{1.3b}$$

where

$$c = \sqrt{a^2+2\sigma^2}. \tag{1.4}$$

Setting $u = T-t$, it can be verified that $A$ and $B$ in (1.2) satisfy the ordinary differential equations (ODEs)

$$A'(u) = \frac{\sigma^2}{2}B(u)^2 - abB(u), \tag{1.5a}$$

$$B'(u) = -aB(u)+1, \tag{1.5b}$$

while $A$ and $B$ in (1.3) satisfy

$$A'(u) = -abB(u), \tag{1.6a}$$

$$B'(u) = -\frac{\sigma^2}{2}B(u)^2 - aB(u)+1. \tag{1.6b}$$

As mentioned above, we will calibrate against the yield curve, which is done as follows. First we choose a particular model for the short rate

$$\mathrm{d}r_t = \mu(t, r_t\,;\alpha)\mathrm{d}t + \sigma(t, r_t;\alpha)\mathrm{d}W_t,$$

with vector parameter $\alpha$ whose associated term structure equation can be solved analytically for $F^T$. In this paper we consider the Vasicek and CIR models. (In fact, we can allow $\alpha$ to depend on $t$ but the inversion of the yield curve is in general difficult to implement, especially in the extended CIR case.) Hence the theoretical term structure is expressed as $\{p(t,T;\alpha) = F^T(t, r_t;\alpha) : 0 \le t \le T, T \ge 0\}$. We also collect bond market price data today to obtain the empirical term structure $\{p^*(0,T){:}T \ge 0\}$ at $t=0$. Then we choose $\alpha$ in such a way that the theoretical curve fits the empirical curve as well as possible, according to some criterion, and which we write as

$$p^*(0,\cdot) \doteq p(0,\cdot\,;\alpha). \tag{1.7}$$

If we are able to do this, then we insert $\alpha$ into $\mu$ and $\sigma$ to pin down the correct martingale measure, and then go on to compute the prices of other interest rate derivatives. The crucial step therefore is how to define $\doteq$ in (1.7), and we will address this point in a moment.

Model calibration approaches in the literature can be divided into the following broad categories. These include (i) descriptive models for the yield curve, (ii) general parametric models, (iii) the least-squares approaches, (iv) maximum likelihood and full Bayesian methods, and (v) splines. Descriptions of these techniques are detailed in chapter 12 of Cairns (2004) and references therein. An encyclopedic account of the calibration methods mentioned here can also be found in James and Webber (2000). In all of these techniques, zero-coupon prices, yield rate curves, or forward rate curves are treated as inputs. Then the model parameters are recovered. We emphasize that fitting models to historical data via several standard statistical techniques, within the context of fixed income research, is a different modeling perspective. Such a perspective is used in other financial modeling endeavors, but not for pricing, where estimated values of parameters must be under a risk-neutral measure. Thus, our inputs are current market prices of bonds.

Let us now go back to (1.7). If we define $\doteq$ pointwise, i.e. if we assume that

$$p^*(0,T) = p(0,T\,;\alpha) \quad \text{for all } T \ge 0,$$

then we have an overdetermined system since $\alpha$ is a finite-dimensional vector. Another natural way of defining $\doteq$ is to choose $\alpha$ in such a way that, for each $T \ge 0$, the mean square error,

$$\int_0^T [p^*(0,u) - p(0,u\,;\alpha)]^2 \mathrm{d}u, \tag{1.8}$$

is minimized. However, it is unlikely that we can analytically find $\alpha$ that will minimize this error since $\alpha$ appears in $F^T$ in a highly nonlinear fashion. Instead, we propose to define $\doteq$ as follows, after rewriting (1.7) as

$$\log p^*(0,\cdot) \doteq A(\cdot\,;\alpha) - r_0 B(\cdot\,;\alpha)$$

for convenience. Multiplying this equation by $u^n$ where $n$ is a non-negative integer, and integrating with respect to $u$ over $[0,T]$, we may define $\doteq$ as

$$\begin{aligned}\int_0^T u^n \log p^*(0,u)\mathrm{d}u = &\int_0^T u^n A(u\,;\alpha)\mathrm{d}u \\ &- r_0\int_0^T u^n B(u\,;\alpha)\mathrm{d}u,\end{aligned} \tag{1.9}$$

for each $T \ge 0$ and for all $n = 0, 1, \ldots$ . In other words, we are requiring that all of the moments of $\log p^*(0,\cdot)$ and $A(\cdot\,;\alpha) - r_0 B(\cdot\,;\alpha)$ be equal. Alternatively, we can interpret (1.9) as equating the weighted averages of the logarithms of the theoretical and empirical bond prices over the interval $[0,T]$.

The outline of this paper is as follows. In section 2 we show how the definition in (1.9) allows us to generate a recursion relation from a particular generating function that involves the vector parameter $\alpha$. The Vasicek and CIR cases will be discussed separately since the functions $A$ and $B$ for each are very different, although the underlying idea that we employ for both cases is the same. In section 3 we present the results of our numerical simulations. We conclude with a brief discussion in section 4.

## 2. Method of generating functions

Assume that the short rate follows the Vasicek or CIR dynamics. For notational simplicity we suppress writing $\alpha$ in the respective $A$ and $B$ functions. Let $s \geq 0$ and define the generating function

$$G_n(s) = \int_0^T u^n \mathrm{e}^{-sB(u)} \mathrm{d}u. \tag{2.1}$$

We choose $B$ in this definition since from (1.5a) and (1.6a) we see that $A$ can be expressed in terms of $B$. It is straightforward to see that

$$\begin{aligned} G_n'(s) &= -\int_0^T u^n \mathrm{e}^{-sB(u)} B(u) \mathrm{d}u, \\ G_n''(s) &= \int_0^T u^n \mathrm{e}^{-sB(u)} B(u)^2 \mathrm{d}u, \end{aligned} \tag{2.2}$$

and, in particular,

$$\begin{aligned} G_n'(0) &= -\int_0^T u^n B(u) \mathrm{d}u, \\ G_n''(0) &= \int_0^T u^n B(u)^2 \mathrm{d}u. \end{aligned} \tag{2.3}$$

Let us also define the sequence

$$q_n = \int_0^T u^n \log p^*(0, u) \mathrm{d}u. \tag{2.4}$$

Then (1.9) becomes

$$q_n = \int_0^T u^n A(u) \mathrm{d}u + r_0 G_n'(0). \tag{2.5}$$

The left-hand side of (2.5) is essentially known since it depends on the empirical term structure at $t = 0$. What we need to do is to express (2.5) in the form $q_n = \Phi_n(\alpha)$ for some function $\Phi_n$ and thus obtain a countably infinite system of equations indexed by $n$. To find $\alpha$ we take as many equations as the dimension of $\alpha$ and solve this system of equations simultaneously.

### *2.1. Vasicek case*

Integrating by parts and using (1.5a) and (2.3), we obtain

$$\begin{aligned} \int_0^T u^n A(u) \mathrm{d}u &= \frac{T^{n+1}}{n+1} A(T) - \frac{\sigma^2}{2(n+1)} \int_0^T u^{n+1} B(u)^2 \mathrm{d}u \\ &\quad + \frac{ab}{n+1} \int_0^T u^{n+1} B(u) \mathrm{d}u \\ &= \frac{T^{n+1}}{n+1} A(T) - \frac{\sigma^2}{2(n+1)} G_{n+1}''(0) \\ &\quad - \frac{ab}{n+1} G_{n+1}'(0). \end{aligned}$$

Then (2.5) simplifies to

$$\begin{aligned} q_n = \frac{T^{n+1}}{n+1} A(T) - \frac{\sigma^2}{2(n+1)} G_{n+1}''(0) - \frac{ab}{n+1} G_{n+1}'(0) \\ + r_0 G_n'(0). \end{aligned}$$

To find $G_n'(0)$ and $G_n''(0)$ we consider the ODE for $B$. Multiplying (1.5b) by $u^n \mathrm{e}^{-sB(u)}$ and integrating with respect to $u$ over $[0, T]$, we have

$$\begin{aligned} \int_0^T u^n \mathrm{e}^{-sB(u)} B'(u) \mathrm{d}u = -a \int_0^T u^n \mathrm{e}^{-sB(u)} B(u) \mathrm{d}u \\ + \int_0^T u^n e^{-sB(u)} \mathrm{d}u \end{aligned}$$

or

$$\int_0^T u^n \mathrm{e}^{-sB(u)} B'(u) \mathrm{d}u = a G_n'(s) + G_n(s)$$

from (2.2). Thus, we can formulate an initial-value problem for the linear first-order ODE

$$\begin{aligned} G_n'(s) + \frac{1}{a} G_n(s) &= H_n(s), \\ H_n(s) &= \frac{1}{a} \int_0^T u^n \mathrm{e}^{-sB(u)} B'(u) \mathrm{d}u, \end{aligned} \tag{2.6}$$

where from (2.1) the initial value is

$$G_n(0) = \frac{T^{n+1}}{n+1}. \tag{2.7}$$

The solution of (2.6) and (2.7) is

$$G_n(s) = \mathrm{e}^{-s/a} \left[ \int_0^s \mathrm{e}^{v/a} H_n(v) \mathrm{d}v + \frac{T^{n+1}}{n+1} \right]. \tag{2.8}$$

Hence,

$$G_n'(s) = H_n(s) - \frac{1}{a} \mathrm{e}^{-s/a} \left[ \int_0^s \mathrm{e}^{v/a} H_n(v) \mathrm{d}v + \frac{T^{n+1}}{n+1} \right]$$

and

$$\begin{aligned} G_n''(s) = H_n'(s) - \frac{1}{a} H_n(s) \\ + \frac{1}{a^2} \mathrm{e}^{-s/a} \left[ \int_0^s \mathrm{e}^{v/a} H_n(v) \mathrm{d}v + \frac{T^{n+1}}{n+1} \right]. \end{aligned}$$

Evaluating these two derivatives at $s = 0$ and using (2.6) and (2.3) gives

$$\begin{aligned} G'_n(0) &= H_n(0) - \frac{T^{n+1}}{a(n+1)} \\ &= \frac{1}{a}\int_0^T u^n B'(u)\mathrm{d}u - \frac{T^{n+1}}{a(n+1)} \\ &= \frac{1}{a}\left[T^n B(T) + nG'_{n-1}(0) - \frac{T^{n+1}}{n+1}\right] \end{aligned}$$

and

$$\begin{aligned} G''_n(0) &= H'_n(0) - \frac{1}{a}H_n(0) + \frac{T^{n+1}}{a^2(n+1)} \\ &= -\frac{1}{a}\int_0^T u^n B(u)B'(u)\mathrm{d}u - \frac{1}{a^2}\int_0^T u^n B'(u)\mathrm{d}u + \frac{T^{n+1}}{a^2(n+1)} \\ &= -\frac{1}{2a}[T^n B(T)^2 - nG''_{n-1}(0)] \\ &\quad - \frac{1}{a^2}\left[T^n B(T) + nG'_{n-1}(0) - \frac{T^{n+1}}{n+1}\right]. \end{aligned}$$

Summarizing, the recursive equations for the Vasicek case are the following:

$$\begin{aligned} q_n &= \frac{T^{n+1}}{n+1}A(T;a,b,\sigma^2) - \frac{\sigma^2}{2(n+1)}G''_{n+1}(0;a) \\ &\quad - \frac{ab}{n+1}G'_{n+1}(0;a) + r_0 G'_n(0;a), \\ G'_n(0;a) &= \frac{1}{a}\left[T^n B(T;a) + nG'_{n-1}(0;a) - \frac{T^{n+1}}{n+1}\right], \\ G''_n(0;a) &= -\frac{1}{2a}[T^n B(T;a)^2 - nG''_{n-1}(0;a)] \\ &\quad - \frac{1}{a^2}\left[T^n B(T;a) + nG'_{n-1}(0;a) - \frac{T^{n+1}}{n+1}\right]. \end{aligned} \tag{2.9}$$

### *2.2. CIR case*

Integrating by parts and using (1.6a) and (2.3), we obtain

$$\begin{aligned} \int_0^T u^n A(u)\mathrm{d}u &= \frac{T^{n+1}}{n+1}A(T) + \frac{ab}{n+1}\int_0^T u^{n+1}B(u)\mathrm{d}u \\ &= \frac{T^{n+1}}{n+1}A(T) - \frac{ab}{n+1}G'_{n+1}(0). \end{aligned}$$

Then (2.5) becomes

$$q_n = \frac{T^{n+1}}{n+1}A(T) - \frac{ab}{n+1}G'_{n+1}(0) + r_0 G'_n(0).$$

Note that in this recursive formula we only need $G'_n(0)$, unlike in the Vasicek case where we also needed $G''_n(0)$.

Multiplying (1.6b) by $u^n \mathrm{e}^{-sB(u)}$ and integrating with respect to $u$ over $[0,T]$, we obtain

$$\begin{aligned} \int_0^T u^n \mathrm{e}^{-sB(u)}B'(u)\mathrm{d}u &= -\frac{\sigma^2}{2}\int_0^T u^n \mathrm{e}^{-sB(u)}B(u)^2\mathrm{d}u \\ &\quad - a\int_0^T u^n \mathrm{e}^{-sB(u)}B(u)du + \int_0^T u^n \mathrm{e}^{-sB(u)}\mathrm{d}u \end{aligned}$$

or

$$\int_0^T u^n \mathrm{e}^{-sB(u)}B'(u)\mathrm{d}u = -\frac{\sigma^2}{2}G''_n(s) + aG'_n(s) + G_n(s)$$

from (2.2). Hence, this time around we have a linear second-order ODE

$$\begin{aligned} G''_n(s) - \frac{2a}{\sigma^2}G'_n(s) - \frac{2}{\sigma^2}G_n(s) &= H_n(s), \\ H_n(s) &= -\frac{2}{\sigma^2}\int_0^T u^n \mathrm{e}^{-sB(u)}B'(u)\mathrm{d}u. \end{aligned} \tag{2.10}$$

This is due to the appearance of the quadratic term in the ODE for $B$. We therefore need two auxiliary conditions. The appropriate boundary conditions can be deduced from (2.1), namely

$$G_n(0) = \frac{T^{n+1}}{n+1}, \qquad G_n(\infty) = 0, \tag{2.11}$$

observing that $B$ is a positive function.

Using the method of variation of parameters, we see that the general solution of (2.10) is

$$\begin{aligned} G_n(s) &= \mathrm{e}^{r_1 s}\left[c_1 - \frac{1}{r_2 - r_1}\int_0^s \mathrm{e}^{-r_1 v}H_n(v)\mathrm{d}v\right] \\ &\quad + \mathrm{e}^{r_2 s}\left[c_2 + \frac{1}{r_2 - r_1}\int_0^s \mathrm{e}^{-r_2 v}H_n(v)\mathrm{d}v\right], \end{aligned} \tag{2.12}$$

where

$$\begin{aligned} r_1 &= \frac{a + \sqrt{a^2 + 2\sigma^2}}{\sigma^2} = \frac{2}{c-a} > 0, \\ r_2 &= \frac{a - \sqrt{a^2 + 2\sigma^2}}{\sigma^2} = -\frac{2}{c+a} < 0 \end{aligned}$$

from (1.4). Here $c_1$ and $c_2$ are arbitrary real numbers that are to be determined so that (2.12) satisfies the boundary conditions in (2.11).

Let

$$\begin{aligned} c_1 &= \frac{2}{\sigma^2(r_1 - r_2)}\int_0^T \frac{u^n B'(u)}{r_1 + B(u)}\mathrm{d}u, \\ c_2 &= \frac{T^{n+1}}{n+1} - c_1. \end{aligned}$$

Then we see that the first condition in (2.11) is automatically satisfied by (2.12). To check the second condition in (2.11) we proceed as follows. First we compute

$$\int_0^s \mathrm{e}^{-r_1 v} H_n(v)\mathrm{d}v = -\frac{2}{\sigma^2}\int_0^s \int_0^T \mathrm{e}^{-r_1 v} u^n \mathrm{e}^{-vB(u)} B'(u)\mathrm{d}u\,\mathrm{d}v$$
$$= -\frac{2}{\sigma^2}\int_0^T u^n B'(u)\int_0^s \mathrm{e}^{-(r_1+B(u))v}\mathrm{d}v\,\mathrm{d}u$$
$$= -\frac{2}{\sigma^2}\int_0^T \frac{u^n B'(u)}{r_1+B(u)}[1-\mathrm{e}^{-(r_1+B(u))s}]\mathrm{d}u.$$

A similar argument gives

$$\int_0^s \mathrm{e}^{-r_2 v} H_n(v)\mathrm{d}v = -\frac{2}{\sigma^2}\int_0^T \frac{u^n B'(u)}{r_2+B(u)}[1-\mathrm{e}^{-(r_2+B(u))s}]\mathrm{d}u.$$

Thus, the first equation in (2.13) can be rewritten as

$$c_1 = \frac{2}{\sigma^2(r_1-r_2)}\lim_{s\to\infty}\int_0^T \frac{u^n B'(u)}{r_1+B(u)}[1-\mathrm{e}^{-(r_1+B(u))s}]\mathrm{d}u$$
$$= \frac{1}{r_2-r_1}\int_0^\infty \mathrm{e}^{-r_1 v} H_n(v)\mathrm{d}v.$$

Furthermore, we can express (2.12) as $G_n(s) = I_1(s) + I_2(s)$, where

$$I_1(s) = \frac{\mathrm{e}^{r_1 s}}{r_2-r_1}\int_s^\infty \mathrm{e}^{-r_1 v} H_n(v)\mathrm{d}v$$

and

$$I_2(s) = c_2\mathrm{e}^{r_2 s} + \frac{2}{\sigma^2(r_1-r_2)}\int_0^T \frac{u^n B'(u)}{r_2+B(u)}[\mathrm{e}^{r_2 s}-\mathrm{e}^{-sB(u)}]\mathrm{d}u.$$

It is clear that $\lim_{s\to\infty} I_2(s) = 0$. An application of L'Hôpital's Rule shows that

$$\lim_{s\to\infty} I_1(s) = \frac{1}{r_2-r_1}\lim_{s\to\infty}\frac{\int_s^\infty \mathrm{e}^{-r_1 v}H_n(v)\mathrm{d}v}{\mathrm{e}^{-r_1 s}}$$
$$= \frac{1}{r_1(r_2-r_1)}\lim_{s\to\infty} H_n(s) = 0,$$

so that $G_n(\infty) = 0$ and the second boundary condition in (2.11) is also satisfied.

So far we have an expression for $G_n(s)$ for any $s \geq 0$; next we compute $G_n'(0)$. Differentiating (2.12) with respect to $s$, we have

$$G_n'(s) = -\frac{1}{r_2-r_1}H_n(s) + r_1\mathrm{e}^{r_1 s}\left[c_1 - \frac{1}{r_2-r_1}\int_0^s \mathrm{e}^{-r_1 v}H_n(v)dv\right]$$
$$+\frac{1}{r_2-r_1}H_n(s) + r_2\mathrm{e}^{r_2 s}\left[c_2 + \frac{1}{r_2-r_1}\int_0^s \mathrm{e}^{-r_2 v}H_n(v)\mathrm{d}v\right],$$

and (2.13) gives

$$G_n'(0) = r_1 c_1 + r_2 c_2 = (r_1-r_2)c_1 + \frac{r_2 T^{n+1}}{n+1}$$
$$= \frac{2}{\sigma^2}\int_0^T \frac{u^n B'(u)}{r_1+B(u)}\mathrm{d}u + \frac{r_2 T^{n+1}}{n+1}.$$

Then

$$G_0'(0) = \frac{2}{\sigma^2}\log\frac{r_1+B(T)}{r_1} + r_2 T,$$

while for $n \geq 1$ we obtain

$$G_n'(0) = \frac{2}{\sigma^2}\left[T^n\log(r_1+B(T)) - n\int_0^T u^{n-1}\log(r_1+B(u))\mathrm{d}u\right]$$
$$+\frac{r_2 T^{n+1}}{n+1}.$$

To simplify the ensuing notation, let us introduce

$$x(u) = -\frac{c-a}{c+a}\mathrm{e}^{-cu}. \tag{2.14}$$

We observe that

$$|x(u)| = \frac{c-a}{c+a}\mathrm{e}^{-cu} \leq \frac{c-a}{c+a} < 1 \tag{2.15}$$

for all $u \geq 0$. This implies that

$$r_1 + B(u) = \frac{2}{c-a} + \frac{2(\mathrm{e}^{cu}-1)}{(c+a)\mathrm{e}^{cu}+(c-a)} = \frac{4c/(c^2-a^2)}{1-x(u)}$$

and

$$\log(r_1+B(u)) = \log\frac{4c}{c^2-a^2} - \log(1-x(u)).$$

Then

$$G_0'(0) = \frac{4}{c^2-a^2}\log\frac{2c/(c+a)}{1-x(T)} - \frac{2T}{c+a}.$$

Moreover, for all $n \geq 1$,

$$\int_0^T u^{n-1}\log(r_1+B(u))\mathrm{d}u = -\int_0^T u^{n-1}\log(1-x(u))\mathrm{d}u$$
$$+\frac{T^n}{n}\log\frac{4c}{c^2-a^2},$$

whence

$$G_n'(0) = \frac{4}{c^2 - a^2}\left[-T^n \log(1 - x(T)) + n\int_0^T u^{n-1}\log(1 - x(u))\mathrm{d}u\right] - \frac{2T^{n+1}}{(c+a)(n+1)}.$$

It remains to evaluate the integral above to complete the specification of $G_n'(0)$.

Before we proceed, let us mention a special function and its properties which we shall need in the sequel. Recall the polylogarithm function (Lewin 1981)

$$\mathrm{Li}_m(z) = \sum_{j=1}^{\infty} \frac{z^j}{z^m},$$

where $|z| < 1$ and $m$ is a positive integer. Note that $\mathrm{Li}_1(z) = -\log(1 - z)$. It can be shown that

$$\begin{aligned}\mathrm{Li}_2(z) &= -\int_0^z \frac{\log(1-v)}{v}\mathrm{d}v,\\ \mathrm{Li}_{m+1}(z) &= \int_0^z \frac{\mathrm{Li}_m(v)}{v}\mathrm{d}v.\end{aligned} \tag{2.16}$$

We are now ready to evaluate $\int_0^T u^{n-1}\log(1 - x(u))\mathrm{d}u$. Suppose first that $n = 1$. Setting $v = x(u)$, we see from (2.15) and (2.16) that

$$\begin{aligned}\int_0^T \log(1 - x(u))\,\mathrm{d}u &= -\frac{1}{c}\int_{x(0)}^{x(T)} \frac{\log(1-v)}{v}\mathrm{d}v\\ &= \frac{1}{c}\mathrm{Li}_2 x(T) - \frac{1}{c}\mathrm{Li}_2 x(0),\end{aligned}$$

thus

$$G_1'(0) = \frac{4}{c^2 - a^2}\left[-T\log(1 - x(T)) + \frac{1}{c}\mathrm{Li}_2 x(T) - \frac{1}{c}\mathrm{Li}_2 x(0)\right] - \frac{T^2}{c+a}.$$

Now suppose that $n \geq 2$. Recall the series

$$\log(1+z) = \sum_{j=1}^{\infty} \frac{(-1)^{j+1}}{j} z^j,$$

which converges for $|z| < 1$. If $z = -x(u)$, then (2.15) gives $|z| < 1$ and

$$\int_0^T u^{n-1}\log(1 - x(u))\mathrm{d}u = \sum_{j=1}^{\infty}\frac{(-1)^{j+1}}{j}\int_0^T u^{n-1}(-x(u))^j\mathrm{d}u. \tag{2.17}$$

But

$$\int_0^T u^{n-1}(-x(u))^j\mathrm{d}u = \left(\frac{c-a}{c+a}\right)^j \int_0^T u^{n-1}\mathrm{e}^{-jcu}\mathrm{d}u$$

and

$$\begin{aligned}\mathrm{e}^{jcu}\int u^{n-1}\mathrm{e}^{-jcu}\mathrm{d}u &= \frac{u^{n-1}}{(-jc)} + \sum_{k=1}^{n-1}\frac{(-1)^k(n-1)(n-2)\cdots(n-k)}{(-jc)^{k+1}}u^{n-k-1}\\ &= -\left[\frac{u^{n-1}}{jc} + \sum_{k=1}^{n-1}\frac{(n-1)!}{(n-k-1)!}\frac{u^{n-k-1}}{(jc)^{k+1}}\right]\\ &= -\left[\frac{u^{n-1}}{jc} + \frac{(n-1)!}{(jc)^n} + \sum_{k=1}^{n-2}\frac{(n-1)!}{(n-k-1)!}\frac{u^{n-k-1}}{(jc)^{k+1}}\right].\end{aligned}$$

Evaluating at $u = 0$ and $u = T$ yields

$$\int_0^T u^{n-1}\mathrm{e}^{-jcu}\mathrm{d}u = -\mathrm{e}^{-jcT}\left[\frac{T^{n-1}}{jc} + \sum_{k=1}^{n-1}\frac{(n-1)!}{(n-k-1)!}\frac{T^{n-k-1}}{(jc)^{k+1}}\right] + \frac{(n-1)!}{(jc)^n}.$$

Hence, from (2.17) we obtain

$$\begin{aligned}&\int_0^T u^{n-1}\log(1 - x(u))\mathrm{d}u\\ &= \sum_{j=1}^{\infty}\frac{x(T)^j}{j}\left[\frac{T^{n-1}}{jc} + \sum_{k=1}^{n-1}\frac{(n-1)!}{(n-k-1)!}\frac{T^{n-k-1}}{(jc)^{k+1}}\right] - \frac{(n-1)!}{c^n}\sum_{j=1}^{\infty}\frac{x(0)^j}{j^{n+1}}\\ &= \frac{T^{n-1}}{c}\mathrm{Li}_2 x(T) - \frac{(n-1)!}{c^n}\mathrm{Li}_{n+1}x(0) + (n-1)!\sum_{k=1}^{n-1}\frac{T^{n-k-1}}{c^{k+1}(n-k-1)!}\mathrm{Li}_{k+2}x(T).\end{aligned}$$

Therefore, for all $n \geq 2$, we conclude that

$$G_n'(0) = \frac{4}{c^2 - a^2}\left[-T^n\log(1 - x(T)) + \frac{nT^{n-1}}{c}\mathrm{Li}_2 x(T) - \frac{n!}{c^n}\mathrm{Li}_{n+1}x(0)\right] + \frac{4n!}{c^2 - a^2}\sum_{k=1}^{n-1}\frac{T^{n-k-1}}{c^{k+1}(n-k-1)!}\mathrm{Li}_{k+2}x(T) - \frac{2T^{n+1}}{(c+a)(n+1)}.$$

Summarizing, the recursive equations for the CIR case are the following:

$$
\begin{aligned}
q_n &= \frac{T^{n+1}}{n+1}A(T;a,b,c) - \frac{ab}{n+1}G'_{n+1}(0;a,c) \\
&\quad + r_0 G'_n(0;a,c), \\
G'_0(0;a,c) &= \frac{4}{c^2-a^2}\log\frac{2c/(c+a)}{1-x(T;a,c)} - \frac{2T}{c+a}, \\
G'_1(0;a,c) &= -\frac{4T}{c^2-a^2}\log(1-x(T;a,c)) - \frac{T^2}{c+a} \\
&\quad + \frac{4}{c^2-a^2}\left[\frac{1}{c}\mathrm{Li}_2 x(T;a,c) - \frac{1}{c}\mathrm{Li}_2 x(0;a,c)\right], \\
G'_n(0;a,c) &= -\frac{4T^n}{c^2-a^2}\log(1-x(T;a,c)) - \frac{2T^{n+1}}{(c+a)(n+1)} \\
&\quad + \frac{4}{c^2-a^2}\left[\frac{nT^{n-1}}{c}\mathrm{Li}_2 x(T;a,c) - \frac{n!}{c^n}\mathrm{Li}_{n+1} x(0;a,c)\right] \\
&\quad + \frac{4n!}{c^2-a^2}\sum_{k=1}^{n-1}\frac{T^{n-k-1}}{c^{k+1}(n-k-1)!}\mathrm{Li}_{k+2} x(T;a,c) \quad (n \geq 2),
\end{aligned} \tag{2.18}
$$

where $x$ is given by (2.14).

## 3. Numerical results

Here we present the results of our numerical simulations in Scilab. In practice, we do not have observed bond prices for all $T \geq 0$ but only for a few maturity dates. In addition, although we can use actual bond market prices, the accuracy of our calibration method should not be tested this way since the Vasicek and CIR short rate models are only approximations of reality. Therefore, in order to test the accuracy of our proposed method, we will use the exact bond price formula to generate the bond prices, i.e. we assume values for $r_0$, $T$, $a$, $b$, and $\sigma$, as well as a finite number of maturity dates (or, equivalently, times to maturity since we are taking $t = 0$ in $u = T - t$), and then evaluate (1.1) at each of these times. More specifically, for the times to maturity we take $u = 1/12$, 1/4, 1/2, 1, 2, 5, 10, 15, and 20 years.

For the Vasicek case it is convenient to treat $\alpha = (a, b, \sigma^2)$ as the unknown vector parameter. We take

$$r_0 = 0.03, \quad T = 20, \quad a = 3.5, \quad b = 0.03, \quad \sigma^2 = 0.3,$$

and compute the corresponding bond prices using (1.1), where $A$ and $B$ are given by (1.2). Including the condition that $p^*(0,0) = 1$, we use splines to generate an interpolating function $p^*(0,u)$ for all $0 \leq u \leq T$ (see figure 1). Using (2.9) with $n = 0$, 1, and 2, we obtain a system of three nonlinear equations in the unknowns $a$, $b$, and $\sigma^2$. Solving this system gives the estimated values

$$\hat{a} = 3.5255, \qquad \hat{b} = 0.0303, \qquad \hat{\sigma}^2 = 0.3110,$$

which agree closely with the actual values used above.

On the other hand, for the CIR case it is convenient to treat $\alpha = (a, b, c)$ as the unknown vector parameter. We take

$$r_0 = 0.03, \quad T = 20, \quad a = 3.2, \quad b = 0.03, \quad c = 2.2.$$

We again compute the corresponding bond prices using (1.1), but with $A$ and $B$ as in (1.3). The corresponding interpolating function is shown in figure 2. Using (2.18) with $n = 0$, 1, and 2, we obtain a system of three nonlinear equations in the unknowns $a$, $b$, and $c$. Solving this system gives the estimated values

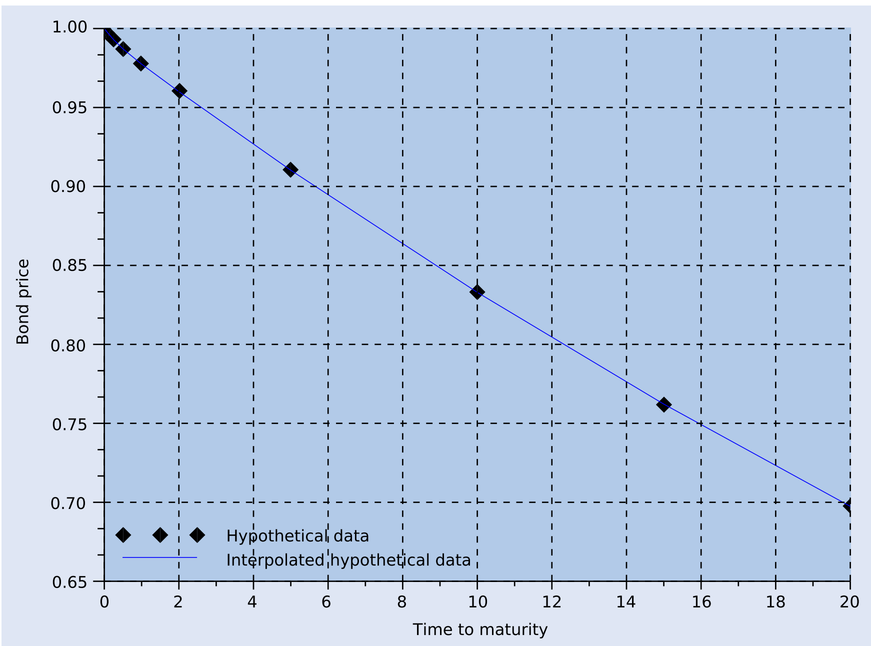


Figure 1. Hypothetical bond prices and interpolating function for the Vasicek case with $r_0 = 0.03$, $T = 20$, $a = 3.5$, $b = 0.03$, and $\sigma^2 = 0.3$.

Table 1. Results.

| $T$ | $\hat{a}$ | $\hat{b}$ | $\hat{\sigma^2}$ |
|---|---|---|---|
| 1/12 | 2.9329 | 0.0300 | 0.2871 |
| 1/4 | 3.4612 | 0.0300 | 0.2980 |
| 1/2 | 3.5545 | 0.0300 | 0.3080 |
| 1 | 3.5520 | 0.0302 | 0.3132 |
| 2 | 3.4292 | 0.0300 | 0.2821 |
| 5 | 2.3528 | 0.0194 | 0.0188 |
| 10 | 3.4154 | 0.0241 | 0.1467 |
| 15 | 3.4732 | 0.0283 | 0.2550 |
| 20 | 3.4940 | 0.0307 | 0.3090 |

$$\hat{a} = 3.2072, \qquad \hat{b} = 0.03, \qquad \hat{c} = 2.2021,$$

again with good agreement with the original parameters. The formulation of our approach enables us to verify if our method will produce 'stable' results as follows. For definiteness, let us consider the Vasicek model and fix $\alpha = (a, b, \sigma^2)$. For different values of the maturity $T$, we would like to see if our method will recover the same values for $a$, $b$, and $\sigma^2$. We choose

$$r_0 = 0.03, \qquad a = 3.5, \qquad b = 0.03, \qquad \sigma^2 = 0.3,$$

as before, but let $T$ vary, i.e. take $T = 1/12$, 1/4, 1/2, 1, 2, 5, 10, 15, and 20.

At each $T$, we take 10 bond prices and estimate the integral in $q_n$. Note that, at each $T$, we generate an interpolating function for all $u \in [0, T]$ and estimate the parameters. The interpolating function would vary as $T$ varies. Then we obtain the following results shown in table 1.

We also remark that we can test the robustness of our method by adding a uniformly distributed random perturbation to the exact bond prices before interpolating and calculating the integral $q_n$. We also obtain similar results as above provided the perturbation is really small. Otherwise, the interpolated function will exhibit 'wiggles' that will greatly affect the estimations if the perturbations are not so small. One possible way around this is not to use spline interpolation but fit an exponential affine interpolating function (using a linear least-squares method to the logarithm of the bond prices).

## 4. Discussion and conclusion

In this article, we have proposed an alternative calibration method based on generating functions to back out the parameters in the Vasicek and CIR models. The definition of the generating function (2.1) can be viewed as a type of integral transform that is applied to the differential equations in (1.5b) and (1.6b). Most integral transforms, such as the Laplace or Fourier transforms, are applicable only to linear equations. However, (1.6b) is a nonlinear equation and yet the transformed equation (2.2) is linear. Thus, (2.1) is an example of a nonlinear integral transform.

Another interesting observation is that the recursive equations (2.9) and (2.18) were obtained as necessary conditions when assuming the Vasicek and CIR models, i.e. if we assume that the short rate follows the Vasicek or CIR dynamics, then necessarily the bond price should satisfy (2.9) and (2.18). Hence, any calibration method for the Vasicek and CIR models that utilizes the yield curve must satisfy these recursive equations, thus serving as a benchmark for such a method.

In the numerical simulations we fixed the short rate $r_0$ throughout, since while this value cannot be observed directly, it can be proxied in principle. Nevertheless, if for some reason $r_0$ is not known or is not reliable, then we can

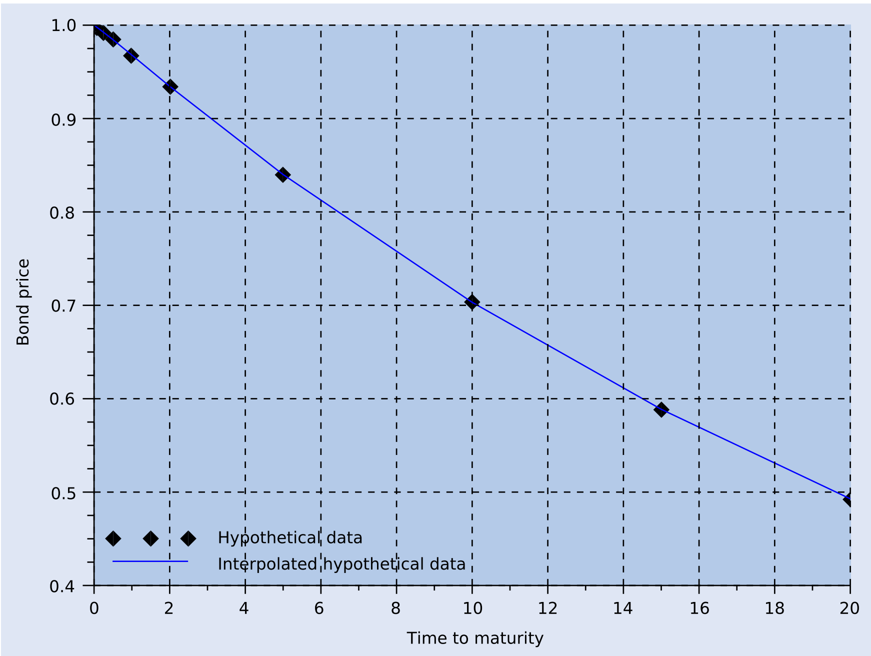


Figure 2. Hypothetical bond prices and interpolating function for the CIR case, where $r_0 = 0.03$, $T = 20$, $a = 3.2$, $b = 0.03$, and $c = 2.2$.

include $r_0$ as another parameter to be calibrated, i.e. we let the empirical term structure $p^*(0,\cdot)$ today tell us what $r_0$ should be. In the implementation we take four values of $n$, instead of three, to generate four linearly independent equations for the vector parameter $\alpha = (a, b, \sigma, r_0)$. We remark that if we were to use a least-squares approach, then we would have to recalculate all the formulas for the minimizing $\alpha$ (if in fact this can be done at all).

As pointed out by Cairns (2004), weighted least-squares methods can also be utilized in model calibration. In some applications the weights are equal, while in others the weights diminish to zero as the term of the fixed-income instrument approaches zero. Our approach also incorporates this desirable feature of varying weights attached to the observed bond market price data; see (2.4), for instance.

From the statistical point of view, our paper presented what could be viewed as the so-called 'generating-function estimators' for the Vasicek and CIR bond price models via the respective recursive equations (2.9) and (2.18). It would therefore be worthwhile to consider and develop statistical techniques that could stabilize the parameters calculated using our approach through time as part of future research directions. This involves the construction of robust estimators or procedures in conjunction with calibration. Robustification is essential to support an approach meant to perform well even if there is some contamination in the data or if the model assumptions used in formulating the procedure are violated by the available data.

## References

Bacinello, A., Ortu, F. and Stucchi, P., Valuation of sinking-fund bonds in the Vasicek and CIR frameworks. *Appl. Math. Finance*, 1996, **3**, 269–394.

Biffis, E., Affine processes for dynamic mortality and actuarial valuations. *Insurance: Math. Econ*., 2004, **37**, 443–468.

Björk, T., *Arbitrage Theory in Continuous Time*, 3rd ed., 2009 (Oxford University Press: Oxford).

Cairns, A., *Interest Rate Models: An Introduction*, 2004 (Princeton University Press: Princeton, NJ).

Cox, J., Ingersoll, J. and Ross, S., A theory of the term structure of interest rates. *Econometrica*, 1985, **53**, 385–408.

Duffie, D. and Kan, R., A yield-factor model of interest rates. *Math. Finance*, 1996, **6**, 379–406.

Elliott, R. and van der Hoek, J., Stochastic flows and the forward measure. *Finance Stochast*., 2001, **5**, 511–525.

Jalen, L. and Mamon, R., Valuation of contingent claims with mortality and interest rate risks. *Math. Comput. Model.*, 2009, **49**, 1893–1904.

James, J. and Webber, N., *Interest Rate Modelling*, 2000 (Wiley: Chichester).

Lewin, L., *Polylogarithms and Associated Functions*, 1981 (Elsevier: New York).

Vasicek, O., An equilibrium characterization of the term structure. *J. Financ. Econ.*, 1977, **5**, 177–188.

Copyright of Quantitative Finance is the property of Routledge and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.