JACKKNIFING BOND OPTION PRICES

BY

PETER C. B. PHILLIPS

Jackknifing Bond Option Prices

Peter C. B. Phillips
Cowles Foundation for Research in Economics, Yale University,
University of Auckland and University of York

Jun Yu
University of Auckland and Singapore Management University

Prices of interest rate derivative securities depend crucially on the mean reversion parameters of the underlying diffusions. These parameters are subject to estimation bias when standard methods are used. The estimation bias can be substantial even in very large samples and much more serious than the discretization bias, and it translates into a bias in pricing bond options and other derivative securities that is important in practical work. This article proposes a very general and computationally inexpensive method of bias reduction that is based on Quenouille's (1956; Biometrika, 43, 353–360) jackknife. We show how the method can be applied directly to the options price itself as well as the coefficients in the models. We investigate its performance in a Monte Carlo study. Empirical applications to U.S. dollar swap rates highlight the differences between bond and option prices implied by the jackknife procedure and those implied by the standard approach. These differences are large and suggest that bias reduction in pricing options is important in practical applications.

For more than three decades continuous time models have proved to be versatile and productive tools in finance. Sundaresan (2000) provides a recent extensive survey of these models and their many financial applications. The models are especially useful with respect to pricing derivative securities where both closed form solutions and numerical methods are used in practical work. This article is concerned with pricing interest rate derivative securities, a practical issue that has been addressed in a number of different ways in the past. One of the oldest and most important approaches is based on modeling the dynamics of factors such as the instantaneous interest rate. According to this approach, to calculate prices

of derivative securities it is necessary to estimate the system parameters that occur in the continuous time specification of the underlying asset. Since only discrete time observations are available, a common practice is to discretize the continuous time system and estimate the resulting discretized model [see, e.g., Chan *et al.*, (1992)]. Unless the exact discrete model is known, as it is in certain special cases [e.g., Phillips (1972)], discretization generally introduces an estimation bias since the internal dynamics between sampling points are ignored. Misspecification bias results in inconsistent estimation [see, Merton (1980), Lo (1988), Melino (1994)] of the parameters of the continuous system with consequent bias effects on derivative prices.

   To circumvent the problem of inconsistent estimation of continuous systems, methods have been proposed to estimate continuous time specifications directly. Among the techniques that have been proposed, the maximum likelihood (ML) approach is naturally appealing in view of its good asymptotic properties in general regular estimation problems, and maximum likelihood estimation (MLE) has become something of a gold standard to aim for in the estimation of continuous time systems. In consequence, many articles have suggested ways of constructing or approximating the likelihood function of a continuous system analytically and of computing it by numerical or simulation methods. Examples of this approach include Lo (1988), Pedersen (1995), Kessler (1995), Durham and Gallant (2002), Aït-Sahalia (1999, 2002, 2003), Brandt and Santa-Clara (2002), and Duffie, Pedersen, and Singleton (2003), to mention only a few. The ultimate goal of all these methods is to approach the MLE in the hope that this estimator will deliver best performance characteristics.

   In spite of its generally good asymptotic properties, the MLE can have a substantial finite sample bias in dynamic models of the type used in financial econometric applications. The estimation bias can be much more substantial than the discretization bias. It is well known to occur in simple models like the first order autoregression (AR) [Hurvicz (1950), Orcutt (1948)] and to be especially acute when autoregressive roots are near unity, corresponding to diffusions with mean reversion parameters close to zero. It is further exacerbated when intercepts and trends are fitted [Orcutt and Winokur (1969), Andrews (1993), Andrews and Chen (1994)]. These bias problems in discrete time dynamic models are manifest in the estimation of continuous time systems, such as diffusion models for short-term interest rates [Ball and Torous (1996), Chapman and Pearson (2000), Yu and Phillips (2001)] and they persist even when the sample size is quite large, as is often the case in financial applications. Similar arguments apply to other commonly used estimation methods in dynamic models, including the general class of extremum estimators.

   The problem of estimation bias is of great importance in the practical use of econometric estimates in asset pricing. The prices of bond options

and other derivative securities hinge crucially on the value of unknown parameters. Of particular importance in diffusion models are the parameters governing volatility and drift. When these parameters are estimated with bias, as occurs with the MLE and many other estimation procedures, estimation bias is transmitted to the pricing formulae for bonds, bond options, and other derivative securities. For instance, when the true mean reversion parameter is 0.1 and 600 weekly observations (i.e., just over 10 years of data) are available to estimate a one-factor square-root diffusion model [Cox, Ingersoll, and Ross (1985)], the bias in the ML estimator of the mean reversion parameter is 391.2% in an upward direction. This estimation bias, together with estimation errors and nonlinearity, produces a 60.6% downward bias in the option price of a discount bond and 2.48% downward bias in the discount bond price. The latter figures are comparable in magnitude to the estimates of bias effects discussed in Hull (2000, Chapter 21.7). The biases would be even larger when less observations are available and do not disappear even with the use of long spans of data, which are currently available. For example, when the true mean reversion parameter is 0.1 and 600 monthly observations (i.e., 50 years of data) are available to estimate the square-root diffusion model, the bias in the ML estimator of the mean reversion parameter is 84.5% in an upward direction. This estimation bias implies a 24.4% downward bias in the option price of a discount bond and a 1.0% downward bias in the discount bond price. Of course, these numbers depend on other aspects of the specification, including the nature of the bond and the maturity of the option, discussed in Sections 1 and 2. The existence of bias in stock option pricing has been noticed in the literature. For example, it is well known that the Black–Scholes stock option price estimates are biased, even when an unbiased volatility estimate is used [Butler and Schachter (1986), Knight and Satchell (1997)]. However, compared with the documented bias effect in stock option pricing where the bias arises entirely from the estimation errors and nonlinearity, the bias effect in bond option pricing is found to be much more dramatic in this article.

To address the problem of biased estimation in continuous time models with its consequential effects on bond option prices, this article introduces bias reduction techniques based on the jackknife [Quenouille (1956)]. While jackknife methods have been extensively used in discrete time models [e.g., Efron (1982), Shao and Tu (1995)], we know of no earlier implementation in continuous time model estimation.

The jackknife has several properties that make it appealing in the present application. The first advantage is its generality. Unlike other bias reduction methods, such as those based on corrections obtained by estimating higher order terms in an asymptotic expansion of the bias, the jackknife technique does not rely on the explicit form of an asymptotic expansion. This means that it is applicable in a broad range of model

specifications and it is not necessary to develop explicit higher order representations of the bias. In the present context, we can, for instance, apply the jackknife technique directly to the quantity of interest (like the option price itself) and hence also reduce some of the bias arising from the estimation errors; see Lo (2003) for an application of the jackknife in the context of the Black–Scholes model. Given the complicated form of options price representations in terms of the underlying process and its parameters, this advantage is significant and makes the method very suitable for empirical implementation. In fact, it is observed that direct use of the jackknife to the options price, provides significant gains relative to bias reduction in the parameters of the continuous time model. Moreover, other methods of parameter bias reduction in dynamic models, like median unbiased procedures [e.g., Andrews (1993)] are only applicable to parameter estimation and are not directly applicable to more complex quantities like options prices, which depend on many other aspects (including distributional details) of the model. A second advantage of the jackknife is that this approach to bias reduction can be used with many different estimation methods, including general methods like MLE. As indicated above, it can also be applied to many different models. This flexibility is illustrated here by implementing the jackknife in both single factor and two-factor affine term structure models both in simulations and in an empirical application. Third, it can be applied in any asset pricing situation (e.g., stock and currency options with stochastic interest rates and interest rate derivatives) where the quantities of interest depend on the estimation of continuous time systems in which finite sample bias arises. Finally and most importantly for practical purposes, unlike many other bias correction methods such as median unbiased estimation, bias function approximation [MacKinnon and Smith (1998)], and bootstrapping, the jackknife is computationally much cheaper to implement. In fact, the method is not more time consuming than the initial estimation itself.

Our findings in this article indicate that the jackknife provides a very substantial improvement in pricing bond options over existing methods. To illustrate, Figure 1 compares the distribution of estimates of the option price of a discount bond obtained by using MLE and jackknifed MLE in a one-factor Cox, Ingersoll, and Ross (CIR) model with 600 monthly observations. As is apparent in the figure, the jackknife estimates are much better centered on the true options price and do not show any appreciable increase in variance. In fact, the root mean squared error (RMSE) of the jackknifed estimates is 12.1% smaller than that of MLE while also providing a bias reduction of 11.5% for 600 observations. Section 2 explores this implementation of the jackknife in detail and shows that a carefully designed jackknife method can lead to a lower value of RMSE, so bias reduction is accomplished without compromising the gains by much larger variability.
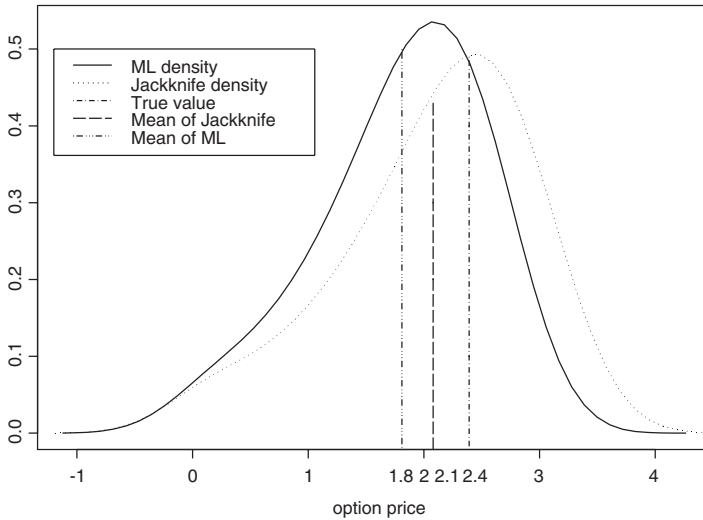
Figure 1
Distribution of jackknife and ML estimates of bond option price
The graphs in the figure show the kernel density of jackknife and ML estimates of bond option prices based on 600 monthly observations and 1000 replications.

In pricing bond options and interest rate derivatives, model specification is known to be important. For instance, Chan *et al.* (1992) show that use of the constant elasticity of variance (CEV) model leads to significant changes in bond option prices compared with alternative models such as the CIR or Vasicek model. Concern over specification has also led to the introduction of more flexible methods of estimation, such as the semi-parametric treatment of diffusion in Aït-Sahalia (1996a) and the fully nonparametric approaches of Stanton (1997) and Bandi and Phillips (2003) that allow users to be agnostic regarding functional form. At least in models where the drift is linear and parametric, the discrete time equivalent model that is satisfied by equispaced observations has the same general autoregressive form, so that dynamic estimation bias of the type discussed above can be expected in all conventional approaches. In consequence, it may be expected that a bias reduction procedure such as the jackknife may be useful even in situations where the continuous time model is misspecified by incorrect specification of the diffusion. Our findings indicate that the jackknife indeed continues to deliver bias reduction in both mean reversion parameter estimation and in pricing bond options under model misspecification. In fact, the results suggest that bias reduction may be more important in practice than correct specification of the diffusion term in pricing bond options.

The article is organized as follows. Using simulated data, Section 1 shows the bias effects of MLE on system parameters, prices of discount

bonds and options on a discount bond in the context of a single factor diffusion model. Section 2 introduces a generic version of the jackknife and shows how it can be implemented in parameter estimation, and bond and option valuation. The simulation performance of these jackknife estimates is compared with that of the ML approach. We also discuss bias and variance trade-offs, consider a version of the jackknife that reduces variability, examine the performance of the jackknife when the model is misspecified, and compare the performance of the jackknife with median unbiased estimation as an alternative method of bias reduction. To demonstrate the usefulness of the jackknife in a more complex and realistic set-up, we then investigate its performance in the context of a two-factor affine diffusion model. Section 3 shows the practical effects of jackknifing in an empirical application with data on London Interbank Offered Rate (LIBOR) and swap rates. Section 4 concludes and outlines some further applications and implications of the approach.

1. Estimation Bias in Continuous Time Models, Bond Pricing, and Bond Option Pricing

We start our discussion with a brief review of some well-known bias results and bias correction methods for discrete time dynamic models. Most relevant in the present context is the fact that standard procedures like ML and least squares (LS) produce downward biased coefficient estimators in the first order AR. Using analytic techniques, Hurvicz (1950) demonstrated the bias effect in the first order AR model with known intercept. Using Monte Carlo techniques, Orcutt (1948) and Orcutt and Winokur (1969) found that the bias is larger when the intercept is fitted and explained the bias enlargement in terms of the induced correlation between the regressor and the residual that results from a fitted intercept. Andrews (1993) showed that the presence of a time trend in the regression further accentuates the autoregressive bias. In these two cases, the biases do not go to zero as the AR coefficient goes to zero and the biases increase as the AR coefficient goes to unity.

In the context of the AR(1) model with an intercept only, Kendall (1954) showed that, to a first-order approximation,

$$E\left[\hat{\phi}\right] - \phi = -\frac{1 + 3\phi}{T} + O\left(\frac{1}{T^2}\right),\tag{1}$$

where T is the sample size and $\hat{\phi}$ is the ML/LS estimator of the AR coefficient ϕ. A natural bias correction method in this simple setting is

$$\hat{\phi}_K = \hat{\phi} + \frac{1 + 3\hat{\phi}}{T}.\tag{2}$$

In the finance literature, Bekaert, Hodrick, and Marshall (1997) used Kendall's method to correct for bias in testing the expectations hypothesis

of the term structure of interest rates. Although feasible in this simple model, where Equation (1) and various higher order extensions of Equation (1) have long been known [e.g., Shenton and Johnson, (1965)], an undesirable property of the correction method is that it is not directly applicable in more complicated set-ups where asymptotic expansion formulae have not been derived.

As an alternative bias correction method in the AR(1) model with fitted intercept and/or time trend, Andrews (1993) proposed a median unbiased estimator of ϕ. The method relies on knowledge of the exact median function of the estimator. Although the procedure is extended to deal with more general AR(p) models in Andrews and Chen (1994), the estimator is no longer exactly median unbiased and it is not available in more complex models where there are usually additional parameter dependencies in the median function.

Similar bias problems occur in the estimation of continuous time dynamic models. As in discrete time models, the problem is worse when the series are persistent. This phenomenon was documented by Chapman and Pearson (2000), for instance, in the context of the following CEV model [c.f. Chan *et al.* (1992)],

$$dr(t) = \kappa(\mu - r(t))dt + \sigma r^\gamma(t)dB(t), \tag{3}$$

where $r(t)$ represents the instantaneous riskless interest rate, $B(t)$ is a standard Brownian motion, and $\theta = (\kappa, \mu, \gamma, \sigma)$ is the vector of unknown system parameters. In this model, $r(t)$ mean-reverts towards the unconditional mean μ with speed captured by κ. The observed data are recorded discretely at $(0, \Delta, 2\Delta, \ldots, T\Delta)$ in the time interval $[0, T\Delta]$, where Δ is the step in a sequence of discrete observations of $r(t)$. Since $r(t)$ is often recorded as the annualized interest rate, if it is observed monthly (weekly or daily), we have $\Delta = 1/12$ (1/52 or 1/252).

Chapman and Pearson (2000) used weighted least squares (WLS) to estimate κ in a discretized version of Equation (3) for daily interest rates. Their simulation findings confirm that the estimate of κ is upward biased and that the bias is significant even when the sample size is as large as 7500. Using the same CEV model, Yu and Phillips (2001) find that alternative Gaussian methods of estimating Equation (3), such as those proposed by Nowman (1997), substantially overestimate κ for daily, weekly, and monthly frequencies, whereas the biases are generally small for the other parameters. Ball and Torous (1996) used WLS and GMM to estimate the discretized model of a restricted version of Equation (3) with $\gamma = 0.5$ for both weekly and monthly interest rates. They found that although both μ and σ can be estimated accurately, the sampling distribution of estimated κ is substantially biased upward.

These results are not surprising because the CEV model has a discrete time formulation that is very similar to an AR(1) model where there is

unconditional heteroscedasticity and with an autoregressive coefficient that is dependent on κ, that is, $\phi = \exp(-\kappa\Delta)$. Using the first-order approximation, we have $\phi \approx 1 - \kappa\Delta$. Then, because κ is often very close to zero in practical work, the model corresponds in discrete time to a "near unit root" model, for which estimation bias problems are well known. Also, the estimation bias for κ equals that for ϕ multiplied by $1/\Delta$ and so is potentially much larger in practice. Since the prices of bonds and bond options also crucially depend on κ, the upward bias in coefficient estimation translates directly into biased bond and option pricing. This important implication of dynamic model estimation bias is explored below in the context of the well-known CIR model specializations of Equation (3) by Cox, Ingersoll, and Ross (1985). There are several reasons why we use the square-root models to study bias effects. First, the square-root models have closed form expressions for the transition and marginal densities. As a result, we can simulate the discrete observations directly from the continuous time models and hence avoid simulation errors. For the same reason, we can perform exact ML estimation directly on the continuous time models and hence avoid the discretization bias. Second, there are known closed form options price formulae for the square-root models and hence no approximation errors are introduced. In summary, all these closed form solutions enable us to quantify estimation bias effects in the most accurate way and to perform a large scale Monte Carlo study. In the empirical application, however, we will also consider a model for which the closed form solutions are not available.

Setting $\gamma = \frac{1}{2}$ in Equation (3), the CIR model has the form

$$dr(t) = \kappa(\mu - r(t))dt + \sigma r^{1/2}(t)dB(t). \tag{4}$$

Feller (1951) and Cox, Ingersoll, and Ross (1985) show that the transition density of $r(t + \Delta)$ conditional on $r(t)$ is $ce^{-u-v}(v/u)^{q/2}I_q(2(uv)^{1/2})$ and the marginal density of $r(t)$ is $w_1^{w_2}r^{w_2-1}e^{-w_1 r}/\Gamma(w_2)$, where $c = 2\kappa/(\sigma^2(1 - e^{-\kappa\Delta}))$, $u = cr(t)e^{-\kappa\Delta}$, $v = cr(t+\Delta)$, $q = 2\kappa\mu/\sigma^2 - 1$, $w_1 = 2\kappa/\sigma^2$, $w_2 = 2\kappa\mu/\sigma^2$, and $I_q(\cdot)$ is the modified Bessel function of the first kind of order q. The transition density together with the marginal density can be used for simulation purposes as well as for obtaining the full ML estimator of θ $(=(\kappa, \mu, \sigma)')$.

Prices of discount bonds and call options on discount bonds based on the square-root model (4) both have analytic solutions. Define $P(t, s)$ as the price at time t of a discount bond that pays-off \$1 at time s and $C(t, \tau; s, K)$ as the value at time t of a call option on a discount bond of maturity data s and of principal L, with exercise (or strike) price K and expiration date τ $(s > \tau > t)$. (Note that, as distinct from options on stock prices, the moneyness of bond options is determined here by the relative size of K to $L \exp(-(s - t)r(t))$. By assuming the market price of interest rate risk is λ, Cox, Ingersoll, and Ross (1985) derived the expression for
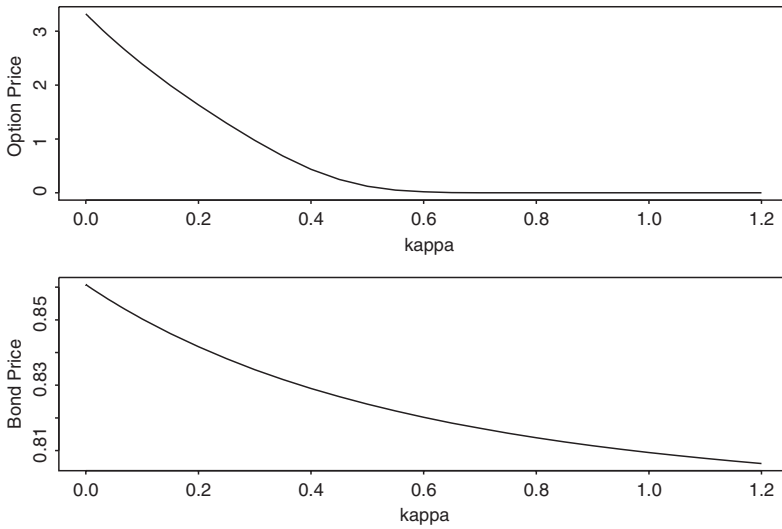
Figure 2
Relationship between κ and option and bond prices
This figure shows the price of a three-year zero coupon bond and the price of one-year European call option on a three-year zero coupon bond as a function of the mean reversion parameter, κ, under a CIR model. The assumed parameters are $\mu = 0.08$, $\sigma = 0.02$. The face value of the coupon bond is $1 and the initial interest rate is 5%. The option on the discount bond has a face value of $100 and a strike price of $87.

$P(t, s)$ and $C(t, \tau; s, K)$. In this article, unless specified explicitly, we assume $\lambda = 0$, which implies that the physical measure is the same as the risk neutral measure.

Both bond and option prices depend on the mean reversion parameter, κ. Figure 2 plots the price of a discount bond and the price of the option on the discount bond as a function of κ. The discount bond is a three-year bond (hence $t = 0$ and $s = 3$) with a face value of $1 and initial interest rate of 5%. The one-year European call option on a three-year discount bond has a face value of $100 and a strike price of $87 (i.e., $t = 0$, $r(t) = 0.05$, $s = 3$, $\tau = 1$, $L = 100$, $K = 87$). These parameters are empirically reasonable and imply that $K/(L\exp(-(s-t)\,r(t))) = 1.011$. Hence, in this case we have to price an out-of-the-money option. We choose $\mu = 0.08$, $\sigma = 0.02$ in model (4). It can be seen that as κ changes both bond and option prices change in a nonlinear and monotonically decreasing fashion. So any bias in estimated κ is transmitted to the corresponding estimates of the bond and option prices. In particular, overestimation of κ leads to underestimation of the bond and option prices. Also, as is apparent in Figure 2, the bond price is much less sensitive to a change in κ than the option price. As a result, we expect bias in κ to have a larger impact on option pricing. Furthermore, in both cases the sensitivity depends on the magnitude of κ. The smaller the κ, the larger the sensitivity. Finally, the nonlinear

Table 1
Finite sample properties of ML and jackknife estimates of κ, bond prices and option prices for weekly data under the one-factor CIR model

	ML estimates					Jackknife estimates				
	κ	μ	σ	Bond	Option	κ	μ	σ	Bond	Option
True	**0.1**	**0.08**	**0.02**	**0.8503**	**2.392**	**0.1**	**0.08**	**0.02**	**0.8503**	**2.392**
Mean	0.4912	0.0802	0.0201	0.8292	0.9426	0.0765	0.0802	0.0200	0.8397	1.2989
SD	0.3847	0.0100	0.0006	0.0174	0.9161	0.5468	0.0102	0.0006	0.0227	1.3567
RMSE	0.5486	0.0100	0.0006	0.0273	1.7148	0.5473	0.0102	0.0006	0.0250	1.7423
True	**0.2**	**0.08**	**0.0283**	**0.8418**	**1.661**	**0.2**	**0.08**	**0.0283**	**0.8418**	**1.661**
Mean	0.5805	0.0803	0.0284	0.8251	0.7163	0.1740	0.0803	0.0283	0.8340	0.9590
SD	0.4048	0.0087	0.0008	0.0162	0.7642	0.5719	0.0091	0.0008	0.0215	1.1264
RMSE	0.5555	0.0087	0.0008	0.0233	1.2150	0.5725	0.0091	0.0008	0.0229	1.3271
True	**0.3**	**0.08**	**0.0346**	**0.8349**	**1.108**	**0.3**	**0.08**	**0.0346**	**0.8349**	**1.108**
Mean	0.6720	0.0803	0.0347	0.8215	0.5422	0.2688	0.0802	0.0347	0.8290	0.6949
SD	0.4171	0.0077	0.0010	0.0151	0.6343	0.5922	0.0082	0.0010	0.0198	0.9160
RMSE	0.5589	0.0077	0.0010	0.0202	0.8502	0.5930	0.0082	0.0010	0.0207	1.0050
True	**0.4**	**0.08**	**0.04**	**0.8291**	**0.7154**	**0.4**	**0.08**	**0.04**	**0.8291**	**0.7154**
Mean	0.7672	0.0803	0.0401	0.8183	0.4100	0.3699	0.0802	0.0401	0.8246	0.5008
SD	0.4329	0.0071	0.0011	0.0141	0.5274	0.6031	0.0076	0.0011	0.0182	0.7429
RMSE	0.5676	0.0071	0.0011	0.0178	0.6095	0.6039	0.0076	0.0011	0.0187	0.7733
True	**0.5**	**0.08**	**0.0447**	**0.8243**	**0.4489**	**0.5**	**0.08**	**0.0447**	**0.8243**	**0.4489**
Mean	0.8648	0.0802	0.0448	0.8154	0.3113	0.4740	0.0802	0.0448	0.8208	0.3612
SD	0.4511	0.0065	0.0013	0.0133	0.4387	0.6101	0.0070	0.0013	0.0167	0.6014
RMSE	0.5801	0.0065	0.0013	0.0160	0.4598	0.6107	0.0070	0.0013	0.0171	0.6078
True	**0.6**	**0.08**	**0.0490**	**0.8203**	**0.2745**	**0.6**	**0.08**	**0.0490**	**0.8203**	**0.2745**
Mean	0.9636	0.0802	0.0491	0.8130	0.2369	0.5788	0.0802	0.0491	0.8175	0.2601
SD	0.4705	0.0061	0.0014	0.0127	0.3648	0.6195	0.0065	0.0014	0.0155	0.4866
RMSE	0.5947	0.0061	0.0014	0.0146	0.3667	0.6199	0.0065	0.0014	0.0158	0.4868

Note: The number of weekly observations is 600 and the number of replications is 1000. SD and RMSE stand for standard deviation and root mean squared error, respectively.

relationships imply that the bias in estimating asset prices can also arise from the estimation errors.

To investigate the bias effect of ML estimation, we choose $\mu = 0.08$, so the mean of the unconditional distribution is fixed at 8%. We also choose a range of values of κ and σ so that the second moment of the unconditional distribution is fixed at 0.025. The range of possible true values of κ is 0.1, 0.2, . . . , 0.6. These are empirically reasonable values and correspond to autoregressive coefficients in the range from 0.9981 to 0.9885 for the weekly frequency and from 0.9917 to 0.9512 for the monthly frequency (see Equation (12)).

Tables 1 and 2 show the means, standard deviations, and RMSEs of the MLEs of κ, μ, σ, and the ML estimated option price and bond price. Table 1 gives the weekly frequency results with sample size 600. This corresponds to just over 10 years of data and represents typical spans used in the empirical literature based on LIBOR and swap data. Although LIBOR data are more useful as far as bond options are concerned, most of the empirical studies of term structure models have been based on Treasury data which typically have a longer span of data. To examine if the

Table 2
Finite sample properties of ML and jackknife estimates of κ, bond prices and option prices for monthly data under the one-factor CIR model

	ML estimates					Jackknife estimates				
	κ	μ	σ	Bond	Option	κ	μ	σ	Bond	Option
True	**0.1**	**0.08**	**0.02**	**0.8503**	**2.392**	**0.1**	**0.08**	**0.02**	**0.8503**	**2.392**
Mean	0.1845	0.0803	0.0201	0.8437	1.8085	0.0933	0.0802	0.0200	0.8492	1.2068
SD	0.1013	0.0069	0.0006	0.0080	0.6920	0.1397	0.0074	0.0006	0.0104	0.9144
RMSE	0.1319	0.0069	0.0006	0.0103	0.9052	0.1399	0.0074	0.0006	0.0105	0.9330
True	**0.2**	**0.08**	**0.0283**	**0.8418**	**1.661**	**0.2**	**0.08**	**0.0283**	**0.8418**	**1.661**
Mean	0.2837	0.0802	0.0284	0.8367	1.2470	0.1975	0.0802	0.0283	0.8408	1.4788
SD	0.1205	0.0053	0.0008	0.0080	0.6297	0.1509	0.0056	0.0008	0.0098	0.8012
RMSE	0.1467	0.0053	0.0008	0.0096	0.7535	0.1509	0.0056	0.0008	0.0099	0.8216
True	**0.3**	**0.08**	**0.0346**	**0.8349**	**1.108**	**0.3**	**0.08**	**0.0346**	**0.8349**	**1.108**
Mean	0.3838	0.0802	0.0347	0.8307	0.8531	0.2997	0.0801	0.0347	0.8340	0.9820
SD	0.1380	0.0044	0.0010	0.0079	0.5324	0.1646	0.0046	0.0010	0.0093	0.6634
RMSE	0.1614	0.0044	0.0010	0.0089	0.5904	0.1646	0.0046	0.0010	0.0093	0.6754
True	**0.4**	**0.08**	**0.04**	**0.8291**	**0.7154**	**0.4**	**0.08**	**0.04**	**0.8291**	**0.7154**
Mean	0.4841	0.0802	0.0401	0.8258	0.5804	0.4007	0.0801	0.0401	0.8284	0.6457
SD	0.1535	0.0039	0.0011	0.0076	0.4310	0.1771	0.0041	0.0011	0.0087	0.5265
RMSE	0.1750	0.0039	0.0011	0.0083	0.4516	0.1771	0.0041	0.0011	0.0087	0.5311
True	**0.5**	**0.08**	**0.0447**	**0.8243**	**0.4489**	**0.5**	**0.08**	**0.0447**	**0.8243**	**0.4489**
Mean	0.5845	0.0802	0.0449	0.8216	0.3926	0.5012	0.0801	0.0448	0.8237	0.4195
SD	0.1676	0.0035	0.0013	0.0073	0.3385	0.1887	0.0036	0.0013	0.0081	0.4050
RMSE	0.1877	0.0035	0.0013	0.0078	0.3431	0.1887	0.0036	0.0013	0.0082	0.4060
True	**0.6**	**0.08**	**0.0490**	**0.8203**	**0.2745**	**0.6**	**0.08**	**0.0490**	**0.8203**	**0.2745**
Mean	0.6851	0.0802	0.0491	0.8181	0.2639	0.6015	0.0801	0.0491	0.8198	0.2687
SD	0.1808	0.0032	0.0014	0.0069	0.2598	0.1999	0.0033	0.0014	0.0076	0.3044
RMSE	0.1998	0.0032	0.0014	0.0073	0.2600	0.1999	0.0033	0.0014	0.0076	0.3045

Note: The number of monthly observations is 600 and the number of replications is 1000. SD and RMSE stand for standard deviation and root mean squared error, respectively.

bias is serious in longest spans of data that are empirically relevant, Table 2 gives results for the monthly frequency with a sample size of 600 which corresponds to 50 years of data. All results are based on 1000 replications.

The following general conclusions emerge from these tables. First, the ML estimator of κ is upward biased and the percentage bias decreases monotonically with the true value of κ. This result is consistent with what is known about dynamic bias in AR/unit root models [e.g., Andrews (1993)], as larger κ corresponds to a smaller AR coefficient. In all cases, the biases are serious for empirically relevant values of κ and sample sizes. In fact we have found that the estimate bias is much more serious than the discretization (either Euler or Milstein) bias for empirically relevant values of κ and sample sizes but are surprised that it has received much less attention than the discretization bias in the continuous time finance literature.[1] Unlike κ, the other parameters in the model are estimated very accurately.

---

[1] The Euler and Milstein discretization bias is obtained based on the same and similar simulation designs, respectively and the results may be obtained from the authors upon request. In general we find that, for empirically relevant parameter settings and sample sizes, the estimation bias is at least 10 times as large as the discretization bias.

Second, although the bias in the ML estimator of the parameter κ is serious, this bias does not translate into a serious bias for the bond price. The outcome is partly explained by Figure 2, where it is clear that the bond price is not very sensitive to changes in κ. However, bonds are always underpriced and this is consistent with the upward bias in estimated κ. In magnitude, the bias monotonically decreases with the true value of κ and stays within the 4% range. By contrast, the options price is substantially underestimated. The percentage bias monotonic decreases in κ and in the worst case considered it is larger than 75%. Hence, bond options are significantly underpriced when κ is estimated by ML.

Finally, the bias of the ML estimator of κ, and the bias in bond and option prices all get smaller as the span increases. This means that on average ML traders would increase the option price when more observations are available. Nonetheless, the bias in κ and the bias in option prices are still nonnegligible even for long spans. These results indicate that biases in the estimation of these quantities must be expected to occur in practical work where the empirical spans are in the same general range as those considered here. The biases are particularly problematic in the case of bond options prices.

It is well known that in the Gaussian AR(1) model (and hence in the Vasicek model defined by Equation (11)), the distribution of the ML estimator of κ is invariant to σ and μ [see, e.g., Andrews (1993)]. While strict invariance results like these do not generally hold in non-Gaussian models, we may continue to expect some degree of scale and mean robustness in estimates of κ because of the linear drift structure. To examine whether the bias of the ML estimator of κ is affected by other parameters, instead of fixing the first two moments of the model, we fix $\mu = 0.08$, $\sigma = 0.02$ but allow κ to vary in model (4). Simulations indicate that the size of the bias in the ML estimator of κ is indeed robust to other parameters in the CIR model.

2. Jackknife Estimation of System Parameters, Bond Prices, and Option Prices

2.1 Jackknife estimation
Quenouille (1956) proposed the jackknife as a solution to finite sample bias in parametric estimation problems. Let T be the number of observations in the whole sample and let the sample be decomposed into m consecutive subsamples each with ℓ observations, so that $T = m \times \ell$. The jackknife estimator of a certain parameter, θ, then utilizes the subsample estimates of θ to assist in the bias reduction process giving

$$\hat{\theta}_{jack} = \frac{m}{m-1}\hat{\theta}_T - \frac{\sum_{i=1}^{m}\hat{\theta}_{li}}{m^2-m}, \tag{5}$$

where $\hat{\theta}_T$ and $\hat{\theta}_{li}$ are the estimates of θ obtained by application of a given method like ML to the whole sample and the ith subsample, respectively. Under quite general conditions which ensure that the bias of the estimates $(\hat{\theta}_T, \hat{\theta}_{li})$ can be expanded asymptotically in a series of increasing powers of T^{-1} it can be shown that the bias in the jackknife estimate $\hat{\theta}_{\text{jack}}$ is of order $O(T^{-2})$ rather than $O(T^{-1})$.

The result can be demonstrated as follows using Sargan's (1976) theorem on the validity of the (Nagar) approximation of the moments of statistical estimator in terms of the moments of the estimator's Taylor expansion as a polynomial of more basic statistics (like sample moments of the data). To fix ideas, suppose $\hat{\theta}_T = \theta_T(p_T)$, where p_T is an N-vector of sample moments of the data with mean μ, whose Taylor development to order k is valid and has the form

$$\theta_{T,k}(p_T) = \sum_{s=0}^{k-1} \frac{1}{s!} \left[\left\{ (p_T - \mu)' \frac{\partial}{\partial p} \right\}^s \theta_T(p) \right]_{p=\mu}. \tag{6}$$

It is frequently the case in practical applications that $p_T - \mu = O_p(T^{-1/2})$ and then Equation (6) produces a corresponding stochastic expansion. Under some mild regularity conditions on the derivatives of $\theta_T(p_T)$ that appear in Equation (6) and the order of magnitude of the moments of $\hat{\theta}_T$ and p, which are assumed to exist, Sargan (1976, Theorems A1 and A2) proved that

$$E\left(|\theta_T(p)|^j\right) = E\left(|\theta_{T,k}(p)|^j\right) + O(T^{-\gamma k}), \quad \gamma > 0, \tag{7}$$

so that for suitably large k, we can replace the jth moment of $\hat{\theta}_T$ by the jth moment of the polynomial approximation $\theta_{T,k}(p)$. This theorem holds rather generally and applies in the present context where $\hat{\theta}_T$ is an econometric estimator of the parameters in the diffusion Equation (4) and p_T is a vector of sample moments of discrete data generated by the model (4). The functional dependence $\hat{\theta}_T = \theta_T(p_T)$ and its Taylor representation (6) may also be obtained indirectly. In the case of extremum estimators like ML, this involves the use of the implicit function theorem and power series inversion of the Taylor expansion of the first-order conditions.

When $\hat{\theta}_T$ is a consistent estimator of θ and when the moment expansion $E(p_T) = \mu + b_1/T + O(T^{-2})$ holds for some constant b_1, we can apply Equation (7) to deduce that for some constant a_1

$$E(\hat{\theta}_T) = \theta + \frac{a_1}{T} + O\left(\frac{1}{T^2}\right) \quad \text{and} \quad E(\hat{\theta}_{\ell i}) = \theta + \frac{a_1}{T/m} + O\left(\frac{1}{(T/m)^2}\right). \tag{8}$$

719

Taking expectations in Equation (5) and substituting the two expressions in Equation (8) leads directly to the expansion

$$E\left(\hat{\theta}_{\text{jack}}\right) = \frac{m}{m-1}\theta + \frac{m}{m-1}\frac{a_1}{T} - \frac{\sum_{i=1}^{m}(\theta + a_1/(T/m))}{m^2 - m} + O\left(\frac{1}{T^2}\right)$$

$$= \theta + O\left(\frac{1}{T^2}\right), \tag{9}$$

reducing the $O(T^{-1})$ bias (8) in the unmodified estimate $\hat{\theta}_T$ to $O(T^{-2})$ in $\hat{\theta}_{jack}$. Note that Equation (9) is invariant to the choice of m to $O(T^{-1})$.

In view of the generality of Equation (7), this bias reduction procedure can be expected to be widely applicable. It is also very easy to implement in practical work. In the present case, we assume that the above theory applies, validating Equations (8) and (9). The quantity θ can be either a parameter (such as κ), a function of parameters (such as the bond option price) or a vector of several such quantities. In the case of bond and options prices, θ will depend on known variables such as t, s and τ, as well as unknown parameters such as κ. These additional dependencies do not affect the validity of the procedure.

In this article we propose to jackknife not only the parameter κ, but also the bond and option prices directly. As some of the bias comes from the estimation errors in the context of term structure models, we have found that there is substantial advantage to the latter procedure of dealing directly with the quantity of interest in implementing the jackknife rather jackknifing the parameter estimates on which the option price depends and plugging this revised estimate into the options price formula. The reason is that the jackknife tends to increase the variance of the quantity being estimated and this additional variance adversely affects the performance of the procedure when the quantity is a very nonlinear function of its arguments, like the option price. In such cases, it appears to be much better to apply the jackknife directly to estimate the option price (see Section 2.4).

In implementing the jackknife (5), it is often convenient to choose $m = 2$ (two subsamples) and this simple choice has very satisfactory performance in bias reduction. In the simulations reported below, we also tried the value $m = 3, 4$ and there are certain advantages to increasing the values of m. In particular, while the mean expansion (9) is invariant to m to the order T^{-1}, the variability of $\hat{\theta}_{\text{jack}}$ depends on m, as is apparent from the following expression for the scaled estimation error of $\hat{\theta}_{\text{jack}}$:

$$\sqrt{T}\left(\hat{\theta}_{\text{jack}} - \theta\right) = \left(1 + \frac{1}{m-1}\right)\sqrt{T}\left(\hat{\theta}_T - \theta\right) - \frac{1}{m-1}\left\{\frac{1}{\sqrt{m}}\sum_{i=1}^{m}\sqrt{\ell}\left(\hat{\theta}_{\ell i} - \theta\right)\right\}.$$

$$\tag{10}$$

In Equation (10) $\sqrt{T}$ scaling is presumed to be appropriate for $\hat{\theta}_T$ and for $\hat{\theta}_{\text{jack}}$, and analogous formulae would apply in the case where there happened to be a faster convergence rate (e.g., due to nonstationarity). It might be anticipated from this expression that larger values of m may help to reduce the variation of $\hat{\theta}_{\text{jack}}$ and, therefore, since Equation (9) still holds, the RMSE of the jackknife estimator $\hat{\theta}_{\text{jack}}$.[2] These heuristics are supported in the present case by the simulation results, which reveal that use of $m = 3, 4$ enables both bias reduction and MSE reduction in estimation.

Since full ML estimation of Equation (4) is feasible for the square-root diffusion, the jackknife procedure can be based on ML. The following specific steps were involved in the implementation of the procedure.[3]

1. Estimate the system parameters by ML using the entire sample.
2. Calculate the bond and option prices based on the ML estimates obtained in Step 1.
3. Estimate the system parameters by ML for each sub-sample.
4. Calculate the bond and option prices based on the ML estimates obtained in Step 3 for each subsample.
5. Calculate the jackknife estimators of κ, and the bond price and option prices using Equation (5).

To compare the performance of the jackknife and ML estimators, we use the same Monte Carlo experiments as the previous section. We first set m to 2. Data are simulated from a one-factor CIR model with the first two moments of the unconditional distribution set at 0.08 and 0.025, and $\kappa = 0.1, 0.2, \ldots, 0.6$. For both weekly and monthly data we choose $T = 600$. The discount bond is a three-year bond with a face value of \$1 and initial interest rate of 5%. The one-year European call option on a three-year discount bond has a face value of \$100 and a strike price of \$87.

Tables 1 and 2 compare the means, standard deviations, and RMSEs of the MLE of κ, μ, σ, and the ML estimated option price and bond price with those obtained by the jackknife in the same experiments. Table 1

---

[2] For example, if $\frac{1}{l} + m/l \rightarrow 0$ as $T \rightarrow \infty$, if the data are weakly dependent, and if the estimates $\hat{\theta}_{li}$ and $\hat{\theta}_T$ are asymptotically normally distributed as

$$\sqrt{l}(\hat{\theta}_{li} - \theta_{li}), \sqrt{T}(\hat{\theta}_T - \theta) \rightarrow_d N(0, \sigma_\theta^2),$$

then, in general,

$$\sqrt{T}(\hat{\theta}_{\text{jack}} - \theta) \rightarrow_d N(0, \sigma_\theta^2),$$

also. However, as m increases, higher order terms in the expansions suggest that the finite sample variation of $\sqrt{T}(\hat{\theta}_{\text{jack}} - \theta)$ decreases with m. A detailed examination of these issues will be provided in later work.

[3] *Matlab* code to implement the procedure in the context of one-factor and two-factor CIR models can be found at http://yoda.eco.auckland.ac.nz/~jyu/research.html. This code covers both the simulations and the empirical work discussed below.

gives the weekly frequency results, while Table 2 gives the monthly frequency results.

First, it is evident that the jackknife successfully reduces the bias in the estimation of κ across all cases. The jackknife works surprisingly well even when the span of data is short. Also, the improvement over ML is greater when the true value of κ is smaller, which is the more relevant case in empirical work. Second, although the ML estimated bond price has only a small downward bias, the jackknife estimated bond price still produces gains in all cases. Third, and most importantly, we find the jackknife estimated bond option price is substantially better than ML. For example, the bias reduction from the jackknife is at least 8% when the weekly frequency was used in the simulation. These gains are of sufficient magnitude to make an important difference in practical work. Fourth, the jackknife method still appears to underprice the option. Fifth, in many cases the bias reductions from the jackknife in pricing options are achieved at the cost of a minor increase in RMSE. Finally, the other two parameters in the model, μ and σ appear to be well estimated by both ML and the jackknife with negligible bias and high precision. For this reason, we will not further report the simulated results on estimated μ and σ.

To understand how the market price of risk affects the performance of the jackknife, we use a similar Monte Carlo design with 300 monthly observations and focus on the empirically more relevant case where $\kappa = 0.1$, but allow λ to take the following empirically reasonable values: $-0.05, -0.07, -0.09$. As a result, the risk neutral measure is different from the physical measure. Table 3 compares the means, standard deviations, and RMSEs of the MLE of κ and the ML estimated option price and bond price. In all cases, the problem of finite sample bias persists for ML and the jackknife still provides an improvement. It is apparent that when the market price of risk becomes larger in absolute value, the estimation bias in the option price is larger in percentage for both ML and the jackknife. However, the effect of the market price of risk on the bond price is less clear.

2.2 Variance reduction

All the above results refer to the case where $m = 2$ and the jackknife is based on only two subsamples. In this case it is apparent from the findings that there is a trade-off between the (often substantial) bias reduction achieved by the jackknife and a marginal increase in the dispersion of the estimates. As argued above, it is possible to reduce the variability of the jackknife estimate (with a small compromise in the bias reduction gains) by using larger values of m. To illustrate the effectiveness of this approach, we use the same Monte Carlo design as before and focus on the empirically more relevant case where $\kappa = 0.1$, considering in all the following four cases: $T = 300, 600$ weekly observations, and $T = 300, 600$ monthly

Table 3
Finite sample properties of ML and jackknife estimates of κ, bond prices and option prices with non-zero price of risk under the one-factor CIR model

λ	Parameter	κ	Bond price	Option price
-0.05	**True value**	**0.1**	**0.8407**	**1.4999**
	ML			
	Mean	0.2762	0.8294	0.7409
	SD	0.1853	0.0119	0.6454
	RMSE	0.2557	0.0164	0.9963
	Jackknife			
	Mean	0.0988	0.8366	0.9742
	SD	0.2606	0.0154	0.9199
	RMSE	0.2606	0.0159	1.0595
-0.07	**True Value**	**0.1**	**0.8367**	**1.0963**
	ML			
	Mean	0.2762	0.8258	0.5050
	SD	0.1853	0.0117	0.5079
	RMSE	0.2557	0.0160	0.7795
	Jackknife			
	Mean	0.0988	0.8326	0.6638
	SD	0.2606	0.0150	0.7247
	RMSE	0.2606	0.0155	0.8439
-0.09	**True value**	**0.1**	**0.8325**	**0.7044**
	ML			
	Mean	0.2762	0.8220	0.3095
	SD	0.1853	0.0114	0.3626
	RMSE	0.2557	0.0155	0.5362
	Jackknife			
	Mean	0.0988	0.8285	0.4030
	SD	0.2606	0.0145	0.5133
	RMSE	0.2606	0.0151	0.5953

Note: The number of monthly observations is 300 and the number of replications is 1000. SD and RMSE stand for standard deviation and root mean squared error, respectively.

observations. Table 4 reports the means, standard deviations, and RMSEs of the jackknife of κ and bond and option prices with $m = 3, 4$ in Equation (5). Compared with the ML results reported in Tables 1–2, it can be seen that the jackknife with $m = 3$ and 4 provides smaller RMSE than ML in all cases and continues to achieve major bias reductions. Comparing the jackknife estimates with different values of m, we find that the jackknife with $m = 4$ provides smallest RMSE in all cases except for the case of 300 weekly observations where $m = 3$ appears to be the optimal choice for pricing the bond and option. This experiment suggests that an optimal m may exist in applications of the jackknife. Further research on this issue is underway and will be reported later.

It is noteworthy that, on average, the jackknife method underestimates κ as well as the option price. At first glance the direction of the bias in the option price estimates seems inconsistent with the direction of the bias in κ. However, since the option price is a nonlinear transformation of κ, an underestimated κ has a different impact in magnitude on the option price from an overestimated κ. As a result, although the jackknife estimate of

Table 4
Finite sample properties of ML and jackknife ($m = 3$ and 4) estimates of κ, bond prices and option prices under the one-factor CIR model

	Parameter	κ	Bond price	Option price
	True value	**0.1**	**0.8503**	**2.3921**
300 weekly observations	Jackknife ($m = 3$)			
	Mean	0.0591	0.8254	0.7261
	SD	0.9198	0.0251	1.0132
	RMSE	0.9207	0.0354	1.9500
	Jackknife ($m = 4$)			
	Mean	0.0656	0.8240	0.6888
	SD	0.8603	0.0242	0.9558
	RMSE	0.8610	0.0357	1.9532
600 weekly observations	Jackknife ($m = 3$)			
	Mean	0.0743	0.8376	1.1890
	SD	0.4846	0.0203	1.1822
	RMSE	0.4853	0.0240	1.6867
	Jackknife ($m = 4$)			
	Mean	0.0728	0.8362	1.1353
	SD	0.4504	0.0195	1.1107
	RMSE	0.4512	0.0240	1.6773
300 monthly observations	Jackknife ($m = 3$)			
	Mean	0.0833	0.8452	1.7256
	SD	0.2299	0.0145	1.1273
	RMSE	0.2306	0.0153	1.3096
	Jackknife ($m = 4$)			
	Mean	0.0854	0.8443	1.6636
	SD	0.2179	0.0138	1.0663
	RMSE	0.2184	0.0151	1.2914
600 monthly observations	Jackknife ($m = 3$)			
	Mean	0.0911	0.8488	2.1403
	SD	0.1205	0.0092	0.8331
	RMSE	0.1208	0.0094	0.8704
	Jackknife ($m = 4$)			
	Mean	0.0901	0.8483	2.0831
	SD	0.1156	0.0089	0.8003
	RMSE	0.1160	0.0091	0.8579

Note: The number of replications is 1000. SD and RMSE stand for standard deviation and root mean squared error, respectively.

κ has little asymmetry (as is apparent in Figure 3), the jackknifed option price becomes asymmetric (as is apparent in Figure 1). Obviously the nonlinearity is the cause of this counter-intuitive result.

2.3 Specification bias versus estimation bias
Specification issues in continuous time modeling of short time interest rates have been a focal point of much recent literature in finance. Important contributions include Chan *et al.* (1992) and Aït-Sahalia (1996a, 1996b), Stanton (1997), Bandi and Phillips (2002), Bandi (2002), and Hong and Li (2005), to mention only a few. Using the CEV model, for example, Chan *et al.* (1992) reject all more restricted nested single factor models and find that the CEV model leads to option prices that are
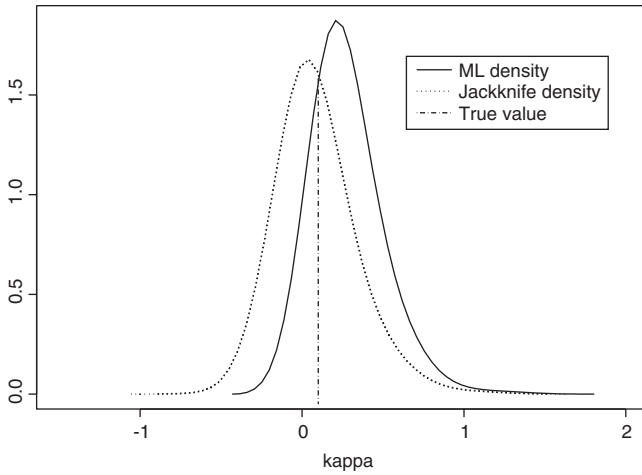
Figure 3
Distribution of ML estimate and jackknife estimate ($m = 4$)
The graphs in the figure show the kernel density of ML estimate and jackknife estimate ($m = 4$) of κ based on 300 monthly observations and 1000 replications.

significantly different from those implied by simpler interest rate processes.

In view of the importance of the diffusion specification to option pricing, it is of interest to compare the magnitude of estimation bias in the drift to bias effects arising from diffusion misspecification. The relative importance of these two effects can be assessed in simulation. To do so, we simulate 600 monthly observations from the following model [Vasicek (1977)],

$$dr(t) = \kappa(\mu - r(t))dt + \sigma dB(t), \tag{11}$$

where $(\kappa, \mu, \sigma)'$ is set at $(0.1, 0.12, 0.015)'$.[4] The discount bond is a three-year bond with a face value of $1 and initial interest rate of 5%. The one-year European call option on a three-year discount bond has a face value of $100 and the strike price such that the ratio of the strike price to the current value of par is 1. The number of replications is 1000. Each simulated sequence is fitted under the (misspecified) CIR model to obtain the ML and jackknife estimates of κ, the bond price, and option price.

Since the diffusion term is misspecified in estimating the CIR model, the ML estimates are biased. However the exact discrete model corresponding

---

[4] We use slightly different parameter values here (specifically, $\mu = 0.12$ and $\sigma = 0.015$ rather than $\mu = 0.08$ and $\sigma = 0.02$) because data from the Vasicek model can become negative and these parameter settings avoid negative values in the 1000 replications used in this comparison.

Table 5
Finite sample properties of ML and jackknife estimates of κ, bond price and option price for the (true) Vasicek model using a (misspecified) fitted CIR model

	Parameter	κ	Bond price	Option price
Strike	**True value**	**0.1**	**0.8371**	**2.2974**
1	ML of CIR			
	Mean	0.1796	0.8226	1.3780
	SD	0.0940	0.0188	1.0482
	RMSE	0.1232	0.0237	1.3901
	Jackknife ($m=2$)			
	Mean	0.0889	0.8348	1.7474
	SD	0.1368	0.0250	1.4649
	RMSE	0.1373	0.0251	1.5625
	Jackknife ($m=4$)			
	Mean	0.0856	0.8328	1.6071
	SD	0.1105	0.0208	1.2325
	RMSE	0.1115	0.0212	1.4095
	ML of Vasicek			
	Mean	0.1890	0.8217	1.4668
	SD	0.1026	0.0202	1.0300
	RMSE	0.1358	0.0255	1.3231

Note: "Strike" is the ratio of the strike price to the present value of the principal. The number of monthly observations is 600 and the number of replications is 1000. SD and RMSE stand for standard deviation and root mean squared error, respectively.

to Equation (11) is [Phillips (1972)]

$$r(t) = \mathrm{e}^{-\kappa\Delta}r(t-\Delta) + \mu\left(1 - \mathrm{e}^{-\kappa\Delta}\right) + \sigma \int_{t-\Delta}^{t} \mathrm{e}^{-\kappa(t-s)}dB(s), \qquad (12)$$

whose autoregressive term is the same as that of the discrete model corresponding to a CIR model. Therefore, we may expect that ML estimates of the drift function in the misspecified model continue to suffer from dynamic estimation bias, making the jackknife desirable. Of course, the ML estimates of the drift function in the correctly specified discrete model (12) will also suffer from dynamic estimation bias. The experimental design in the simulation enables us to isolate the bias arising in the estimation of the drift from that due to misspecification of the diffusion.

Table 5 compares the means, standard deviations, and RMSEs of the ML and jackknife estimates of κ, the bond price, and the option price; true values of these quantities also being shown for comparison purposes. The true bond value is obtained using the analytic formula given in Vasicek (1977) and the true option value is calculated based on the analytic formula derived by Jamshidian (1989). Table 5 reports results for ML estimation of the correctly specified (Vasicek) model obtained from the exact discrete model (12), ML estimation of the CIR model where the diffusion function is misspecified, and jackknife estimates based on the misspecified CIR model.

It is clear from Table 5 that the bias effect plays an important role in all cases. For example, in comparing the ML and jackknife estimates of the

misspecified CIR model when $m = 2$, the jackknife method reduces the bias in the bond price from -1.73% to -0.27%, and the bias in option prices from -40.02% to -23.94%. These are substantial improvements, indicating that the jackknife continues to be a very effective tool of bias reduction even in misspecified situations. When $m = 4$, the jackknife method reduces the bias in the bond price from -1.73% to -0.51%, and the bias in option prices from -40.02% to -30.05%, while also achieving reductions in RMSE over the $m = 2$ setting. Finally, comparing the ML estimates of the correctly specified (Vasicek) model with the jackknife estimates ($m = 4$) of the misspecified CIR model, we find that the jackknife continues to reduce the bias in the bond price, from -1.83% to -0.51%, and the bias in options prices from -36.15% to -29.85%. The improvements indicate that the bias arising from estimating the drift term is generally more serious than that arising from misspecification of the diffusion.[5]

2.4 Jackknife versus median unbiased estimation

In addition to the jackknife, median unbiased estimation (MUE) introduced by Andrews (1993) can also remove bias in coefficient estimation of the discrete AR(1) model. This method relies on full knowledge of the exact median function of the estimator, which can only be obtained by simulation. For continuous time models with nonlinear diffusions the situation is exacerbated because the distribution is non-Gaussian, the median function must be approximated using a computationally highly intensive Monte Carlo method that depends on specific values of the other parameters in the model. In consequence, MUE is not a practically feasible method.

In spite of these practical limitations, the MUE procedure provides a very interesting benchmark for evaluating the success of bias reduction procedures in dynamic models. The MUE is obtained by transforming the ML estimate $\hat{\kappa}$ with the inverse median function m_d^{-1}, where the median function $m_d(\kappa)$ is found by running extensive simulations over a wide range of parameter values κ.[6] In this way the MUE procedure utilizes a great deal of information about the distribution of the ML estimator and, at least when the assumptions underlying the construction of m_d are valid, we can expect no better bias reduction procedure than this.

In the present study, we implemented the MUE in the CIR model, where its performance can be directly compared with that of the jackknife and ML. Using simulated median functions, we correct the bias by

---

[5] In the simulation we also tried to have the CIR and Vasicek models reversed. That is, we simulated data from the CIR model but fitted the Vasicek model to the simulated data. We found that the relative importance of the estimation bias over the misspecification bias is more serious than that reported in the present article.

[6] The median functions may be obtained from the authors upon request.

Table 6
Finite sample properties of ML and jackknife estimates of κ, bond prices and option prices of ML, jackknife and MUE under the one-factor CIR model

	Parameter	κ	Bond price	Option price
	True value	**0.1**	**0.8503**	**2.3921**
300 monthly	ML			
observations	Mean	0.2762	0.8381	1.4071
	SD	0.1853	0.0125	0.8959
	RMSE	0.2557	0.0174	1.3315
	Jackknife ($m=2$)			
	Mean	0.0988	0.8460	1.8186
	SD	0.2606	0.0163	1.2555
	RMSE	0.2606	0.0169	1.3803
	Jackknife ($m=4$)			
	Mean	0.0845	0.8443	1.6636
	SD	0.2179	0.0138	1.0663
	RMSE	0.2184	0.0151	1.2914
	MUE			
	Mean	0.0916	0.8554	2.8015
	SD	0.1953	0.0168	1.4050
	RMSE	0.1954	0.0173	1.4634
	Jackknife plug-in ($m=2$)			
	Mean	0.0988	0.8584	2.8825
	SD	0.2606	0.0564	2.1560
	RMSE	0.2606	0.0570	2.2110
	Jackknife plug-in ($m=4$)			
	Mean	0.0845	0.8562	2.9417
	SD	0.2179	0.0223	1.7644
	RMSE	0.2184	0.0231	1.8480

Note: The number of monthly observations is 300 and the number of replications is 1000. SD and RMSE stand for standard deviation and root mean squared error, respectively.

constructing the median unbiased estimator of κ as in Andrews (1993). In a simulation study, 300 monthly observations are used to estimate the CIR model for parameter values $\kappa=0.1$, $\mu=0.08$, $\sigma=0.02$, and the number of replications is 1000. Table 6 reports the means, standard deviations, and RMSEs of estimates of κ, bond price, and option price from ML, jackknife with $m=2$, 4, and MUE, where the bond and bond option are defined in the same way as before. For comparison purposes, we also plug-in the jackknife parameter estimates to the bond and option price formulae and report the results in Table 6. In terms of the bias reduction in κ, MUE is comparable with the jackknife but has smaller RMSE. The results suggest that MUE is very effective in reducing dynamic bias, as indeed it is designed to do. Not surprisingly, the plug-in MUE works better than the plug-in jackknife for pricing bonds and options as it provides smaller biases as well as smaller RMSEs. However, the plug-in MUE is not superior to the estimates obtained from jackknifing the quantities of interest directly. Although, in magnitude, the bias in the bond and option prices from MUE is comparable with that from jackknifing bond and option prices with $m=2$, the RMSE is larger for MUE, particularly for option prices.

These results confirm that the idea of jackknifing specific quantities of interest directly, rather than plugging in bias reduced parameter estimates is likely to be especially important in practice, as in the present setting where the object is to achieve gains in pricing derivative securities where the formulae are complicated functions of several fundamentals including unknown parameters. It is a characteristic feature of the jackknife method that it permits corrections to be implemented directly on the ultimate quantity of interest. This is an important operational distinction between the jackknife and other bias reduction techniques like MUE, which have more limited applicability in view of distributional and invariance property restrictions.

The reason jackknifing option prices directly works better than various plug-in methods (including the plug-in jackknife) is because of the nonlinear nature of pricing applications. It is already well known that nonlinearity can cause plug-in methods to produce biased option price estimates even when the parameter estimate is unbiased [Butler and Schachater (1986)]. Our results demonstrate that bias effects can be much more significant when the parameter estimate itself is biased and the pricing function is heavily nonlinear. The nonlinearity of the options pricing formula has other important implications in finance. For instance, when comparing alternative option valuation models, Christoffersen and Jacobs (2004) found that a method can often perform well out-of-sample in a dimension that corresponds to the loss function applied in estimating or calibrating the forecasts, but not necessarily well in other dimensions. These differences in option pricing performance are largely driven by the nonlinearities of option pricing functions. By avoiding the use of plug-in operations entirely, appropriate use of the jackknife can be expected to be more robust against the effects of nonlinearity.

2.5 A multifactor model

Bias effects are also present in multifactor models. For example, the finite sample bias in the mean reversion estimators can be found to be substantial in all seven multifactor models in Aït-Sahalia and Kimmel (2003). The bias occurs not only in the approximate ML estimator proposed by Aït-Sahalia (2003) but also in the exact ML estimator.

To demonstrate the usefulness of the jackknife in more complex and realistic frameworks, we consider a two-factor CIR model where there are two independent factors, each evolving over time according to a square-root process:

$$dy_i(t) = \kappa_i(\mu_i - y_i(t))dt + \sigma_i y_i^{1/2}(t)dB_i(t), \quad i = 1, 2. \tag{13}$$

This model is a special case of the more general class of multifactor affine models considered in Duffie and Kan (1996). In the terminology of

Dai and Singleton (2000), it is an A_2 (2) model. This model is chosen in the present study for the following reasons. First, the model has an analytic solution for the prices of discount bonds, so it is not necessary to have to solve some differential equations to obtain a solution, unlike the case of a general affine model. Second, the model has a closed-form likelihood function and therefore lends itself to an exact ML approach.

Define $P(t, s)$ to be the price at time t of a discount bond that pays-off \$1 at time s and we have

$$P(t, s) = A_1(t, s)A_2(t, s)e^{-B_1(t,s)y_1 - B_2(t,s)y_2}, \qquad (14)$$

where

$$A_i(t, \tau) = \left(\frac{2\gamma_i e^{(\kappa_i + \lambda_i + \gamma_i)(\tau - t)/2}}{(\kappa_i + \lambda_i + \gamma_i)(e^{\gamma_i(\tau - t)} - 1) + 2\gamma_i} \right)^{2\kappa_i \mu_i / \sigma_i^2},$$

$$B_i(t, \tau) = \frac{2(e^{\gamma_i(\tau - t)} - 1)}{(\kappa_i + \lambda_i + \gamma_i)(e^{\gamma_i(\tau - t)} - 1) + 2\gamma_i},$$

$$\gamma_i = \sqrt{(\kappa_i + \lambda_i)^2 + 2\sigma_i^2},$$

and λ_i is the market price of interest rate risk associated with factor i.[7]

Let $C(t, \tau; s, K)$ be the value at time t of a call option on a discount bond of maturity data s and of principal L, with exercise (or strike) price K and expiration date τ ($s > \tau > t$). Chen and Scott (1992) give the expression for $C(t, \tau; s, K)$. They show that although $C(t, \tau; s, K)$ does not have an analytic expression, the calculation is reduced to univariate numerical integrations of cummulative density functions (CDFs) and probability density functions (PDFs) of noncentral chi-square variates.

The state vector $y_i(t)$ is not observable to econometricians. Instead, we observe the yields of discount bonds with different maturities. Let $Y(t, s)$ be the bond yield of maturity $s - t$ and Equation (14) implies the following relationship between bond yields and factors:

$$\begin{pmatrix} Y(t, s_1) \\ Y(t, s_2) \end{pmatrix} = \begin{pmatrix} \dfrac{B_1(s_1 - t)}{s_1 - t} & \dfrac{B_2(s_1 - t)}{s_1 - t} \\ \dfrac{B_1(s_2 - t)}{s_2 - t} & \dfrac{B_2(s_2 - t)}{s_2 - t} \end{pmatrix} \begin{pmatrix} y_1(t) \\ y_2(t) \end{pmatrix} - \begin{pmatrix} \dfrac{\ln(A_1(s_1 - t)A_2(s_1 - t))}{s_1 - t} \\ \dfrac{\ln(A_1(s_2 - t)A_2(s_2 - t))}{s_2 - t} \end{pmatrix}.$$

$$(15)$$

It can be seen that a panel of only two yields is needed to back out the state vector and avoid stochastic singularity, a feature that means that

---

[7] See Dai and Singleton (2000) for a discussion of the identification issues of market price of risk.

when the number of yields (say M) used is larger than the number of factors (say N) the model is always rejected [Piazzesi (2002)]. In this case the likelihood function of the state vector is the product of two likelihood functions, each from the univariate CIR model and hence has a closed form expression. As a result, the likelihood function of the observed bond yields is obtained in closed-form by way of standard change of variable techniques.

There exist other techniques to break singularities. For example, one can add measurement errors to the yield equations. One way to do this is to assume all yields are observed with errors. However, this implies that exact ML is no longer applicable. Alternatively, one can assume that among M yields, N yields are observed without error, whereas the other $M - N$ yields involve measurement errors. However, such an assumption seems completely arbitrary. Another difficult problem facing these alternative approaches is how to model the measurement errors. To avoid these controversial issues, we use a panel of two yields (with maturities of six-month and five-year) to estimate the two-factor CIR model (i.e., $M = N = 2$).

To compare the performance of the jackknife and ML estimators in the context of the two-factor CIR model, we design a small scale Monte Carlo experiment. A typical empirical finding in the two-factor CIR model is that one factor has a very slow rate of mean reversion and the other has a considerably faster reverting rate. Judged by the empirical results reported in the literature [see, e.g., Chen and Scott (1993), Geyer and Pichler (1999), Jagannathan, Kaplin and Sun (2003)], a set of parameters is chosen for (κ_1, κ_2): (0.5, 0.01). Other parameter values selected are $\mu_1 = 0.03$, $\mu_2 = 0.02$, $\sigma_1 = 0.001$, $\sigma_2 = 0.001$, $\lambda_1 = 0$, and $\lambda_2 = 0$. The initial values for the factors are 0.03 and 0.02 in order to price bonds and bond options. We minimize the negative of the log-likelihood of the model with respect to $\kappa_1, \kappa_2, \mu_1, \mu_2$, and assume the other parameters are known in order to alleviate the computational burden. Both ML and the jackknife ($m = 2$ and 3) are used to estimate the model. For the three-year discount bond we choose a face value of \$1. The call option on a three-year discount bond has a face value of \$100 and a strike price of \$87. Table 7 reports the results based on 1000 samples of 300 and 600 weekly observations on one-month and five-year zero rates.

First, as in the one-factor CIR model, ML estimates the mean reversion parameters with large upward biases. The bias is 143% for κ_1 and 735% for k_2 when 300 weekly observations are used. The bias is 66% for κ_1 and 519% for k_2 when 600 weekly observations are used. It can be seen that the biases are smaller but remain substantial when the sample size increases. The sizes of these biases generally agree with those in Aït-Sahalia and Kimmel (2003). On the contrary, the jackknife delivers substantial bias reduction in the estimation of κs in both cases and is particularly effective

Table 7
Finite sample properties of ML and jackknife ($m = 3$ and 4) estimates of κ_1, κ_2, bond prices and option prices under the two-factor CIR model

	Parameter	κ_1	κ_2	Bond price	Option price
	True value	**0.5**	**0.01**	**0.8607**	**3.3150**
300 weekly	ML				
observations	Mean	1.2163	0.0835	0.8792	4.7835
	SD	0.7000	0.0019	0.0095	0.7138
	RMSE	1.0015	0.0073	0.0208	1.6328
	Jackknife ($m = 2$)				
	Mean	0.3914	0.0638	0.8729	4.3384
	SD	1.1639	0.0032	0.0148	1.7892
	RMSE	1.1757	0.0061	0.0192	2.0613
	Jackknife ($m = 3$)				
	Mean	0.4610	0.0691	0.8747	4.5084
	SD	0.7453	0.0026	0.0130	0.9968
	RMSE	0.7468	0.0061	0.0191	1.5549
600 weekly	ML				
observations	Mean	0.8310	0.0619	0.8727	4.2917
	SD	0.1290	0.0013	0.0084	0.6643
	RMSE	0.2380	0.0040	0.0146	1.1812
	Jackknife ($m = 2$)				
	Mean	0.4517	0.0395	0.8661	3.7960
	SD	0.2534	0.0025	0.0125	0.9934
	RMSE	0.2557	0.0034	0.0136	1.1037
	Jackknife ($m = 3$)				
	Mean	0.4340	0.0447	0.8674	3.9069
	SD	0.1953	0.0020	0.0110	0.8766
	RMSE	0.1996	0.0032	0.0129	1.0577

Note: The number of replications is 1000. SD and RMSE stand for standard deviation and root mean squared error, respectively.

for κ_1. Second, as expected, the biases in the ML estimators of κs translate into biases for the bond price and more so for the bond option price. For example, when 300 weekly observations are used, the bias is 2.15% in bond price and 44.3% in option price. On the other hand, the jackknife always offers improvements in bond pricing. In the case of 300 weekly observations, the jackknife reduces the bias in the bond price from 2.15% to 1.42% ($m = 2$) and to 1.63% ($m = 3$). In the case of 600 weekly observations, the jackknife reduces the bias in the bond price from 1.39% to 0.63% ($m = 2$) and to 0.78% ($m = 3$). The improvement is more dramatic for option pricing. In the case of 300 weekly observations, the jackknife reduces the bias in the option price from 44.3% to 30.9% ($m = 2$) and to 36.0% ($m = 3$). In the case of 600 weekly observations, the jackknife reduces the bias in the option price from 29.5% to 14.5% ($m = 2$) and to 17.9% ($m = 3$). In most cases, these bias reductions from the jackknife in estimating the mean reversion parameters and pricing options are achieved with a decrease in RMSE. For example, in all cases the jackknife with $m = 3$ provides smaller RMSE than ML. In the case of 600 observations, the jackknife with $m = 2$ provides smaller RMSE than ML in pricing options.

3. Empirical Application

3.1 Data

In the empirical analysis we use weekly U.S. dollar LIBOR and swap rates obtained from Datastream for the period from August 12, 1992 to July 24, 2002. Figure 4 plots the time series of six-month LIBOR rates and the five-year swap rate. The choice of LIBOR data over Treasury data is made because LIBOR rates are more relevant to pricing derivatives [see Dai and Singleton (2000) for a discussion of important differences between LIBOR data and Treasury data]. Using the bootstrap method, we obtain the LIBOR zero curve from all available swap rates and LIBOR rates. We follow Dai and Singleton (2000) by assuming that there is no default risk, but note that the default of risk can be handled in the framework of Duffie and Singleton (1999). The sample size is 520 and chosen to be close to those used in the simulation study to help in calibrating the results with the simulation.

We choose the observed data to be the six-month and five-year zero rates. Table 8 shows the sample sizes, means, standard deviations, first seven autocorrelations, and Phillips' (1987) $Z(t)$ unit root test statistic (with a fitted intercept in the regression) for both series. The presence of a unit root cannot be rejected at the 10% level in both series. These results, together with the form of the sample autocorrelogram, suggest that both
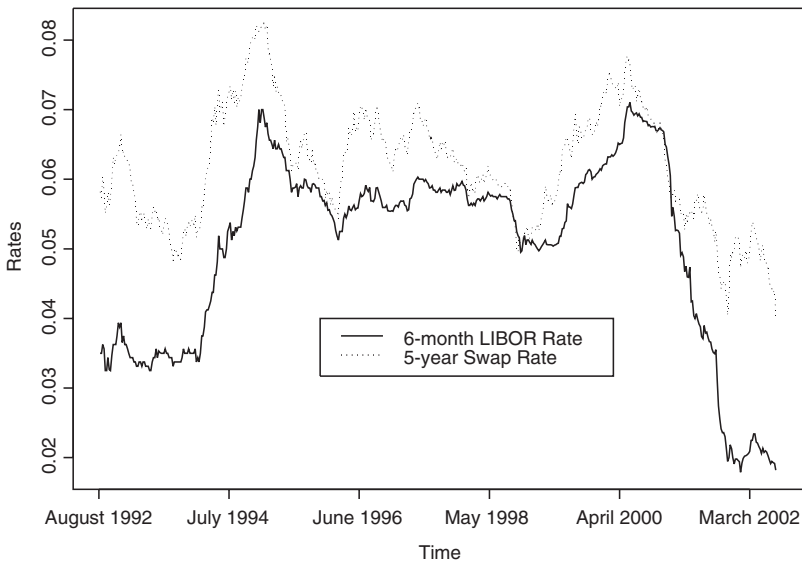

Figure 4
The dynamics of LIBOR and five-year swap rates
The figure shows paths of weekly LIBOR and five-year swap rates from August 12, 1992 to July 24, 2002.

Table 8
Summary statistics and unit root tests for weekly 6-month and 5-year zero rates

Data	6-month	5-year
Number of Observations	520	520
Mean	0.0496	0.0603
Standard Deviation	0.0132	0.00787
Autocorrelation ρ_1	0.991	0.979
Autocorrelation ρ_2	0.982	0.960
Autocorrelation ρ_3	0.972	0.941
Autocorrelation ρ_4	0.961	0.920
Autocorrelation ρ_5	0.949	0.897
Autocorrelation ρ_6	0.937	0.873
Autocorrelation ρ_7	0.925	0.850
$Z(t)$ test	0.0589	−1.205
10% critical value	−2.5699	−2.5699

Note: Data source is Datastream. The LIBOR zero curve from August 12, 1992 to July 24, 2002 and calculated from all available swap rates and LIBOR rates using the bootstrapping method discussed in Hull (2000). $Z(t)$ is the Phillips' Z-statistic and the critical value is obtained from the model with a fitted intercept.

interest rate series are highly persistent. Hence, standard estimation methods for diffusion equations can be expected to lead to significant bias in estimating the correlation and mean reversion coefficients.

3.2 The one-factor model

In this comparison, we focus on a three-year discount bond, a one-year call option, and a half-year call option on a discount bond. In all cases the initial value for the short-term interest rate is set at 5%. For the discount bond we choose a face value of $1. The call option on a three-year discount bond has a face value of $100 and various strike prices so that the ratio of the strike price to the current value of par takes each of the following values: 0.95, 1, and 1.05. These ratios correspond to in-the-money, at-the-money, and out-of-the-money situations, respectively.

To estimate the one-factor CIR model, we assume that the six-month zero rate provides a good proxy to the instantaneous riskless rate. We first estimate the square-root model by ML and jackknife methods and then obtain the ML estimated bond and option prices as well as the jackknife estimates of the two prices.

Table 9 reports the results with $m = 2$. The jackknife estimate of κ is 0.0416 and is 45.1% smaller than the MLE; the jackknife estimate of the bond price is 0.34% higher than its ML counterpart. More importantly, the call option values differ substantially between the two methods. The biggest percentage differences are for the out-of-the-money option. For example, for the half-year option, the price of the out-of-the-money option implied by the jackknife method is 26.5% larger than that obtained from the MLE. All these results are consistent with the magnitudes and directions of the biases and differences between the jackknife and ML estimates that were found in the simulation studies.

Table 9
Empirical estimates of one-factor CIR model

						Option price		
							Exercise price	
							Current value of par	
Method	κ	μ	σ	Bond price	Option Exp. (t: years)	0.95	1	1.05
ML	0.0758	0.0328	0.0324	0.8656	0.5	6.798	2.704	0.199
Jackknife	0.0416	0.0225	0.0324	0.8685		7.077	2.962	0.252
ML					1	8.729	4.634	0.908
Jackknife						8.987	4.890	1.090

Note: We fit the following one-factor CIR model

$$dr(t) = \kappa(\mu - r(t))dt + \sigma r^{1/2}(t)dB(t)$$

to the 6-month zero rate which is assumed to be good proxy to the instantaneous rate.

Table 10
Empirical estimates of two-factor CIR model

Method	κ_1	μ_1	σ_1	λ_1	Option price
ML	0.2170	0.0043	0.024	−1.201	0.1413
Jackknife	0.0865	0.0040	0.020	−.981	0.2111
	κ_2	μ_2	σ_2	λ_2	Bond price
ML	0.0337	0.0248	0.0339	0.1413	0.8071
Jackknife	0.0262	0.0096	0.0355	−0.1125	0.8198

Note: We fit the following two-factor CIR model

$$dy_i(t) = \kappa_i(\mu_i - y_i(t))dt + \sigma_i y_i^{1/2}(t)dB_i(t), \quad i = 1, 2$$

to the 6-month and 5-year zero rates without treating the 6-month rate as an approximation to the instantaneous rate.

3.3 Two-factor models

In this comparison, we consider two multifactor specifications. The first is the two-factor CIR model discussed in Section 2 while the second one is the two-factor central tendency model proposed by Balduzzi, Das, and Foresi (1998) which will be defined below. In both models, we will use six-month and five-year zero rates to estimate parameters and then price a three-year discount bond and a one-year call option on a discount bond. For the discount bond we choose a face value of $1. The call option on a three-year discount bond has a face value of $100 and a strike price of $87.

Table 10 reports the results for the two-factor CIR model when ML and the jackknife with $m = 2$ are used. The initial values for the two factors are set at 4.5% and 0.5%. Several interesting results emerge from this table. First, ML yields the familiar empirical finding about the mean reversion parameters that was mentioned earlier for the two-factor CIR model—that one factor has a very slow rate of mean reversion (0.0337)

and the other has a considerably faster rate (0.2170). However, the jack-knife estimates suggest that both mean reversion parameters are overestimated by ML. The estimated differences between the two methods for the two mean reversion parameters are 22.3% and 60.1%, respectively. According to the jackknife estimate, although one factor still reverts toward its mean faster than the other, the difference between these two mean reversion parameters is 0.0603 and is much smaller than what has been documented in the literature. Indeed both factors are very close to unit root processes according to the jackknife estimates. Second, the jackknife estimate of the bond price is 1.57% higher than its ML counterpart. For the one-year call option, the price implied by the jackknife method is 49.4% larger than that obtained from the MLE.

The last model studied in the application is the central tendency model which, under the physical measure, takes the form of

$$
\begin{aligned}
d\theta(t) &= \beta(\bar{\theta} - \theta(t))dt + \eta\sqrt{\theta(t)}dB_1(t), \\
dr(t) &= \kappa(\theta(t) - r(t))dt + \sigma dB_2(t),
\end{aligned}
\tag{16}
$$

where $r(t)$ is the instantaneous rate, $\theta(t)$ is a long-term mean of $r(t)$, which stochastically follows a CIR process. By assuming that λ_i is the market price of the interest rate risk associated with factor i, this model is an affine $A_1(2)$ model. Despite its popularity in empirical studies, this model does not have a closed-form likelihood function, nor analytical expressions for prices of discount bonds and bond options. To estimate parameters from discrete time observations, we use the approximate maximum likelihood (AML) method developed by Duffie, Pedersen, and Singleton (2003). To price the discount bond, we solve the Ricatti equations (see Equations (6) and (7) in Dai and Singleton, 2000) using Runge-Kutta methods. Finally, the option price is computed by Fourier inversion of a relevant characteristic function, following the suggestion of Duffie, Pan and Singleton (2000).

Table 11 reports the empirical results when AML and the jackknife with $m = 2$ are used. The initial values for the two factors are set at 8% and 4%. Several interesting results emerge from this table. First, the AML estimate of β is very similar to the jackknife estimate of β, suggesting that the reversion parameter of the first factor is estimated with little bias by ML. However, the jackknife estimates suggest that the mean reversion parameter of the second factor (κ) is overestimated by ML. The estimated differences between the two methods is 55.8%. According to the jackknife estimate, therefore, the instantaneous rate reverts very slowly toward its long run mean. Second, the jackknife estimate of the bond price is 1.49% higher than its ML counterpart. For the one-year call option, the price implied by the jackknife method is 48.86% larger than that obtained from the MLE.

Table 11
Empirical estimates of two-factor central tendency model

Method	β	$\bar{\theta}$	η^2	κ	σ^2
ML	0.2501	0.0932	0.0109	0.4631	$0.4653\ e^{-5}$
Jackknife	0.2568	0.0952	0.0137	0.2049	$0.4352\ e^{-5}$
	λ_1	λ_2	Bond price	Option price	
ML	-0.0011	0.0027	0.8362	3.4459	
Jackknife	-0.0015	0.0026	0.8487	5.1297	

Note: We fit the following two-factor central tendency model

$$d\theta(t) = \beta(\bar{\theta} - \theta(t))dt + \eta\sqrt{\theta(t)}dB_1(t),$$
$$dr(t) = \kappa(\theta(t) - r(t))dt + \sigma dB_2(t),$$

to the 6-month and 5-year zero rates without treating the 6-month rate as an approximation to the instantaneous rate.

4. Conclusions and Implications

Bias in the estimation of the parameters of continuous time models by standard methods such as ML translates into bias in pricing bonds and bond options. In cases where the parameters take on realistic values, we have found that these biases can be substantial, particularly in the case of bond options. The procedure we propose here for reducing the bias involves the use of subsample estimates and a version of the jackknife. Simulations show the procedure to be highly effective in the one-factor and two-factor CIR models and to offer substantial improvements in pricing bond options and marginal improvements in pricing bonds over the usual ML approach. The greatest gains are in the substantial bias reductions that the jackknife method provides. But use of multiple sub-samples in the construction of the jackknife enables reductions in both bias and mean squared error, so the gains from bias reductions are not lost in variance increases.

An interesting feature of the proposed method is that it can be used to reduce bias even when the diffusion of the model is misspecified, thereby offering an additional advantage over standard methods. Our simulation findings indicate that the dynamic estimation bias arising from the use of standard estimation methods can be even larger than the specification bias arising from misspecifying the diffusion. Moreover, using the jackknife in a model where the diffusion is misspecified turns out to be less biased than using ML in the correctly specified model.

This article applies the approach to price discount bonds and options in the context of one-factor and two-factor CIR models estimated by ML and two-factor central tendency model estimated by AML. However, the technique itself is quite general and can be applied in many other contexts and models with little modification. For example, the method extends to a broader range of model specifications, including the CEV model

[Chan *et al.* (1992)], extended one-factor models [Hull and White (1990)], two-factor equilibrium models [Brennan and Schwartz (1979, 1982), Longstaff and Schwartz (1992), Langetieg (1980), Countadon (1982)], the semiparametric model [Aït-Sahalia (1996a)] with parametric drift, models with stochastic volatility [Andersen and Lund (1997)], the multi-factor affine family that has been the dominant class of term structure models [e.g., Duffie and Kan (1996), Dai and Singleton (2000)], the affine jump-diffusion models [Duffie, Pan and Singleton (2000), Chacko and Das (2002)], semi-affine models [Duffee (2002), Duarte (2004)], and regime switching model [Bansal and Zhou (2002)]. Because the drift is linear and parametric in all the models, we expect the bias problem to be present in these models also and again our method should be useful.

In more complicated models such as those just mentioned, the analytic form of the likelihood function is often unavailable and so exact ML is infeasible. However, the proposed jackknife method can be used in connection with other estimation methods. Examples include simulated GMM [Duffie and Singleton (1993)] and its refinement EMM [Gallant and Tauchen (1996)], indirect inference [Gourieroux, Monfort and Renault (1993)], continuous time GMM [Hansen and Scheikman (1995)], approximate ML [Aït-Sahalia (2002, 2003), Duffie, Pedersen, and Singleton (2003)], simulated ML [Pedersen (1995), Brandt and Santa-Clara (2002), Durham and Gallant (2002)], and methods via the empirical characteristic function [Singleton (2001), Knight and Yu (2002)]. Finally, many other interest-rate-contingent claims can be treated in a similar way. Examples include coupon-bearing bonds, caps, swaptions, captions, mortgage-back securities, and stock and currency options with stochastic interest rates. Since all these interest-rate-contingent claims are nonlinear functions of the system parameters, any bias in the estimation of the system parameters will carry over to pricing the interest-rate-contingent claims. The situation in these cases is analogous to the one explored here and the proposed jackknife method can be used to reduce the bias in pricing the contingent claims.

The term structure models are estimated using either time series data or small panels in this article. Some recent term structure literature makes use of full panel information [e.g., Ball and Torous (1996), Duffee and Stanton (2002)]. As shown by Ball and Torous and more so by Duffee and Stanton, the finite sample bias problem persists even when considerable cross-sectional information is employed, so there is a continuing need for bias correction procedures. The jackknife method should continue to be useful in such cases and appropriate extensions will be considered in later work.

References

Aït-Sahalia, Y., 1996a, "Nonparametric Pricing of Interest Rate Derivative Securities," *Econometrica*, 64, 527–560.

Aït-Sahalia, Y., 1996b, "Testing Continuous-time Models of Spot Interest Rate Derivative Securities," *Review of Financial Studies*, 9, 385–426.

Aït-Sahalia, Y., 1999, "Transition Densities for Interest Rate and Other Nonlinear Diffusions," *Journal of Finance*, 54, 1361–1395.

Aït-Sahalia, Y., 2002, "Maximum Likelihood Estimation of Discretely Sampled Diffusion: A Closed-form Approximation Approach," *Econometrica*, 70, 223–262.

Aït-Sahalia, Y., 2003, "Closed-form Likelihood Expansion for Multivariate Diffusion," working paper, Princeton University, Princeton, NJ.

Aït-Sahalia, Y., and R. Kimmel, 2003, "Estimating Affine Multifactor Term Structure Models Using Closed-form Likelihood Expansion," working paper, Princeton University, Princeton, NJ.

Andersen, T. G., and J. Lund, 1997, "Estimation Continuous-time Stochastic Volatility Models of the Short-term Interest Rate," *Journal of Econometrics*, 77, 343–377.

Andrews, D. W. K., 1993, "Exactly Median-unbiased Estimation of First Order Autoregressive/unit Root Models," *Econometrica*, 61, 139–166.

Andrews, D. W. K., and H. Chen, 1994, "Approximately Median-unbiased Estimation of Autoregressive Models," *Journal of Business and Economic Statistics*, 12, 187–204.

Balduzzi, P., S. Das, and S. Foresi, 1998, "The Central Tendency: A Second Factor in Bond Yields," *Review of Economics and Statistics*, 80, 62–72.

Ball, C. A., and W. N. Torous, 1996, "Unit Roots and the Estimation of Interest Rate Dynamics," *Journal of Empirical Finance*, 3, 215–238.

Bansal, R., and H. Zhou, 2002, "Term Structure of Interest Rates with Regime Shifts," *Journal of Finance*, 57, 1997–2043.

Bandi, F. M., 2002, "Short Term Interest Rate Dynamics: A Spatial Approach," *Journal of Financial Economics*, 65, 73–110.

Bandi, F. M., and P. C. B. Phillips, 2003, "Fully Nonparametric Estimation of Scalar Diffusion Models," *Econometrica*, 71, 241–283.

Bekaert, G., R. J. Hodrick, and D. A. Marshall, 1997, "On Biases in Tests of the Expectations Hypothesis of the Term Structure of Interest Rates," *Journal of Financial Economics*, 44, 309–348.

Brandt, M. W., and P. Santa-Clara 2002, "Simulated Likelihood Estimation of Diffusions with an Application to Exchange Rate Dynamics in Incomplete Markets," *Journal of Financial Economics*, 30, 161–210.

Brennan, M. J., and E. S. Schwartz, 1979, "A Continuous Time Approach to the Pricing of Bonds," *Journal of Banking and Finance*, 3, 133–155.

Brennan, M. J., and E. S. Schwartz, 1982, "An Equilibrium Model of Bond Pricing and a Test of Market Efficiency," *Journal of Financial and Quantitative Analysis*, 17, 70–100.

Butler, J.S., and B. Schachter, 1986, "Unbiased Estimation of the Black/Scholes Formula," *Journal of Financial Economics*, 15, 341–357.

Countadon, G., 1982, "The Pricing of Options on Default-free Bonds," *Journal of Financial and Quantitative Analysis*, 17, 75–100.

Chacko, G., and S. Das, 2002, "Pricing Interest Rate Derivatives: A General Approach," *Review of Financial Studies*, 15, 195–241.

Chan, K. C., G. A. Karolyi, F. A. Longstaff, and A. B. Sanders 1992, "An Empirical Comparison of Alternative Models of Short Term Interest Rates," *Journal of Finance*, 47, 1209–1227.

Chapman, D. A., and N. Pearson, 2000, "Is the Short Rate Drift Actually Nonlinear?" *Journal of Finance*, 55, 355–388.

Chen, R. R., and L. Scott, 1992, "Pricing Interest Rate Options in a Two-factor Cox–Ingersoll–Ross Models of the Term Structure," *Review of Financial Studies*, 5, 613–636.

Chen, R. R., and L. Scott, 1993, "Maximum Likelihood Estimation of a Multi-factor Equilibrium Model of the Term Structure of Interest Rates," *Journal of Fixed Income*, 3, 14–31.

Christoffersen, P., and K. Jacobs, 2004, "The Importance of the Loss Function in Option Valuation," *Journal of Financial Economics*, 72, 291–318.

Cox, J., Ingersoll, J., and S. Ross, 1985, "A Theory of the Term Structure of Interest Rates," *Econometrica*, 53, 385–407.

Dai, Q., and K. J. Singleton, 2000, "Specification Analysis of Affine Term Structure Models," *Journal of Finance*, 55, 1943–78.

Duarte, J., 2004, "Evaluating an Alternative Risk Preference in Affine Term Structure Models," *Review of Financial Studies*, 17, 379–404.

Duffee, G. R., 2002, "Term Premia and Interest Rate Forecasts in Affine Models," *Journal of Finance*, 57, 405–443.

Duffee, G. R., and R. Stanton, 2002, "Estimation of Dynamic Term Structure Models," working paper, Haas School of Business, University of California, Berkeley.

Duffie, D., and R. Kan, 1996, "A Yield-factor Model of Interest Rate," *Mathematical Finance*, 6, 379–406.

Duffie, D., J. Pan, and K. J. Singleton, 2000, "Transform Analysis and Asset Pricing for Affine Jump-diffusions," *Econometrica*, 68, 1343–1376.

Duffie, D., L. Pedersen, and K. J. Singleton, 2003, "Modelling Sovereign Yield Spreads: A Case Study of Russian Debt," *Journal of Finance*, 58, 119–159.

Duffie, D., and K. J. Singleton, 1993, "Simulated Moments Estimation of Markov Models of Asset Prices," *Econometrica*, 61, 929–952.

Duffie, D., and K. J. Singleton, 1999, "Modeling Term Structures of Defaultable Bonds," *Review of Financial Studies*, 12, 929–952.

Durham, G., and A. R. Gallant, 2002, "Numerical Techniques for Maximum Likelihood Estimation of Continuous-time Diffusion Processes," *Journal of Business and Economic Statistics*, 20, 297–316.

Efron, B., 1982, *The Jackknife, the Bootstrap and Other Resampling Method*, SIAM, Philadelphia.

Feller, W. 1951, "Two Singular Diffusion Problems," *Annals of Mathematics*, 54, 173–182.

Gallant, A. R., and G. Tauchen, 1996, "Which Moments to Match?" *Econometric Theory*, 12, 657–681.

Geyer, A., and S. Pichler, 1999, "A State-space Approach to Estimate and Test Multifactor Cox–Ingersoll–Ross Models of the Term Structure," *Journal of Financial Research*, 22, 107–130.

Gouriéroux, C., A. Monfort, and E. Renault, 1993, "Indirect Inference," *Journal of Applied Econometrics*, 8, S85–S118.

Hansen, L. P., and J. A. Scheinkman 1995, "Back to Future: Generating Moment Implications for Continuous-time Markov Processes," *Econometrica*, 63, 767–804.

Hong, Y., and H. Li, 2005, "Nonparametric Specification Testing for Continuous-Time Models with Applications to Term Structure of Interest Rates," *Review of Financial Studies*, 18, 37–84.

Hull, J., 2000, *Options, Futures, and Other Derivatives* (2d ed.), Prentice-Hall, New Jersey.

Hull, J., and A. White, 1990, "Pricing of Interest-rate-derivative Securities," *Review of Financial Studies*, 3, 573–592.

Hurvicz, L., 1950, "Least Square Bias in Time Series," in T. Koopmans (ed.), *Statistical Inference in Dynamic Economic Models*, Wiley, New York, pp. 365–383.

Jagannathan, R., A. Kaplin, and S. Sun, 2003, "An Evaluation of Multi-factor CIR Models Using LIBOR, Swap Rates, and Cap and Swaption Prices," *Journal of Econoemtrics*, 116, 113–146.

Jamshidian, F., 1989, "An Exact Bond Option Formula," *Journal of Finance*, 44, 205–209.

Kendall, M. G., 1954, "Notes on Bias in the Estimation of Autocorrelation," *Biometrika*, 41, 403–404.

Kessler, M., 1995, "Estimation of an Ergodic Diffusion from Discrete Observation," *Scandinavian Journal of Statistics*, 24, 211–229.

Knight, J. L., S. E. Satchell, 1997, "Existence of Unbiased Estimators of the Black/Scholes Option Price, Other Derivatives, and Hedge Ratios," *Econometric Theory*, 13, 791–807.

Knight, J. L., and J. Yu, 2002, "The Empirical Characteristic Function in Time Series Estimation," *Econometric Theory*, 18, 691–721.

Langetieg, T. C., 1980, "A Multivariate Model of the Term Structure," *Journal of Finance*, 35, 71–97.

Lo, A. W., 1988, "Maximum Likelihood Estimation of Generalized Itô Processes with Discretely Sampled Data," *Econometric Theory*, 4, 231–247.

Lo, Y., 2003, *Bias Estimation of the Black-Scholes Model*, Honours Dissertation, University of Auckland.

Longstaff, F. A., and E. S. Schwartz, 1992, "Interest Rate Volatility and the Term Structure: A Two-factor General Equilibrium Model," *Journal of Finance*, 47, 1259–1282.

MacKinnon, J. G., and A. A. Smith 1998, "Approximate Bias Correction in Econometrics," *Journal of Econometrics*, 85, 205–230.

Melino, A., 1994, "Estimation of Continuous-time Models in Finance," in Sims, C. A. (ed.), *Advances in Econometrics: Sixth World Congress*, Vol. II, Cambridge University Press, Cambridge.

Merton, R. C., 1980, "On Estimating the Expected Return on the Market: An Exploratory Investigation," *Journal of Financial Economics*, 8, 323–361.

Nowman, K. B., 1997, "Gaussian Estimation of Single-factor Continuous Time Models of the Term Structure of Interest Rates," *Journal of Finance*, 52, 1695–1703.

Orcutt, G. H., 1948, "A Study of the Autoregressive Nature of the Time Series Used for Tinbergen's Model of the Economic System of the United States," *Journal of the Royal Statistical Society, Series B*, 10, 1–45.

Orcutt, G. H., and H.S. Winokur, 1969, "First Order Autoregression: Inference, Estimation and Prediction," *Econometrica*, 37, 1–14.

Pedersen, A., 1995, A New Approach to Maximum Likelihood Estimation for Stochastic Differential Equations Based on Discrete Observation," *Scandinavian Journal of Statistics*, 22, 55–71.

Phillips, P. C. B., 1972, "The Structural Estimation of Stochastic Differential Equation Systems," *Econometrica*, 40, 1021–1041.

Phillips, P. C. B., 1987, "Time Series Regression with a Unit Root," *Econometrica*, 55, 277–301.

Piazzesi, M., 2002, "Affine Term Structure Models," forthcoming in *Handbook of Financial Econometrics*.

Quenouille, M. H., 1956, "Notes on Bias in Estimation," *Biometrika*, 43, 353–360.

Sargan, J. D., 1976, "Econometric Estimators and the Edgeworth Approximation," *Econometrica*, 44, 421–448.

Shao, J., and D. Tu, 1995, *The Jackknife and the Bootstrap*, Springer, New York.

Shenton, L. R., and W. Johnson, 1965, "Moments of a Serial Correlation Coefficient," *Journal of the Royal Statistical Society, Series B*, 27, 308–320.

Singleton, K. J., 2001, "Estimation of Affine Asset Pricing Models using the Empirical Characteristic Function," *Journal of Econometrics*, 102, 111–141.

Stanton, R., 1997, "A Nonparametric Model of Term Structure Dynamics and the Market Price of Interest Rate Risk," *Journal of Finance*, 52, 1973–2002.

Sundaresan, S. M., 2000, "Continuous-time Methods in Finance: A Review and an Assessment," *Journal of Finance*, 55, 1569–1622.

Vasicek, O., 1977, "An Equilibrium Characterization of the Term Structure," *Journal of Financial Economics*, 5, 177–186.

Yu, J., and P. C. B. Phillips, 2001, "A Gaussion Approach for Estimating Continuous Time Models of Short Term Interest Rates," *The Econometrics Journal*, 4, 211–225.