Simulation-Based Estimation of Contingent-Claims Prices

Peter C. B. Phillips
Yale University, University of Auckland, University of York, and Singapore
Management University

Jun Yu
Singapore Management University

A new methodology is proposed to estimate theoretical prices of financial contingent claims whose values are dependent on some other underlying financial assets. In the literature, the preferred choice of estimator is usually maximum likelihood (ML). ML has strong asymptotic justification but is not necessarily the best method in finite samples. This paper proposes a simulation-based method. When it is used in connection with ML, it can improve the finite-sample performance of the ML estimator while maintaining its good asymptotic properties. The method is implemented and evaluated here in the Black-Scholes option pricing model and in the Vasicek bond and bond option pricing model. It is especially favored when the bias in ML is large due to strong persistence in the data or strong nonlinearity in pricing functions. Monte Carlo studies show that the proposed procedures achieve bias reductions over ML estimation in pricing contingent claims when ML is biased. The bias reductions are sometimes accompanied by reductions in variance. Empirical applications to U.S. Treasury bills highlight the differences between the bond prices implied by the simulation-based approach and those delivered by ML. Some consequences for the statistical testing of contingent-claim pricing models are discussed. (*JEL* C11, C15, G12)

Pricing financial contingent claims, whose values depend on the price of an underlying asset, has been an important topic in modern financial economics. Some well-known examples include Black-Scholes (1973); Merton (1973); Vasicek (1977); Cox, Ingersoll, and Ross (1985); Heston (1993); Duan (1996); and Duffie, Pan, and Singleton (2000). Often the underlying asset is assumed

to follow a parametric time-series model, commonly formulated in continuous time, and the price of the contingent claim, often known as the theoretical price, is derived by using no-arbitrage arguments. The resulting price of the contingent claim is a function of the parameters in the time-series model. The functional form of this dependence is almost always complicated and nonlinear.

Since the parameters of the underlying asset are usually unknown, they are generally replaced by time-series estimates in the contingent-claim pricing formulas. Consequently, the statistical properties of the theoretical contingent-claim price estimate critically hinge on those of the parameter estimates. For example, the sampling variation in the estimated contingent-claim price depends on the sampling variation in the estimated parameters. Hence, the choice of method for parameter estimation is important and the topic has received a great deal of attention in the literature (see, for example, Boyle and Ananthanarayanan 1977; Ball and Torous 1984; Lo 1986; Chan et al. 1992; Aït-Sahalia 1999).

Perhaps the most direct method for parameter estimation is to use historical time-series data on the underlying asset price. It has often been argued that when the model for the underlying asset is correctly specified, the preferred basis for estimation and inference should be maximum likelihood (ML) (see, for example, Ball and Torous 1984; Lo 1986; Aït-Sahalia 1999, 2002). There are strong reasons for this choice. Primary among these is the fact that the ML estimator (MLE) has desirable asymptotic properties of consistency, normality, and efficiency under broad conditions (Huber 1967) in stationary time-series settings. Moreover, when the MLE is used in pricing formulas, one naturally expects the good asymptotic properties of the MLE to transfer over to the corresponding contingent-claim price. The theoretical price of a contingent claim is a smooth nonlinear function of the system parameters being estimated, so that plug-in estimates of contingent-claim prices are themselves MLEs in view of the invariance property of maximum likelihood (e.g., Zehna 1966). In consequence, these plug-in pricing estimates have all the desirable asymptotic properties of the MLE. Of course, ML is a very general tool of estimation and inference so that it has wide applicability in this context and, at least for stationary time series, its good asymptotic properties are well established. The ML approach therefore provides a convenient framework for estimation and inference in asset pricing models (cf. Lo 1986).

Despite its generally good asymptotic properties, ML is not necessarily the best estimation method for contingent-claim prices in finite samples for three reasons. First, since the price of a contingent claim is a nonlinear transformation of the system parameters, insertion of even unbiased estimators into the pricing formulas will not assure unbiased estimation of a contingent-claim price (Ingersoll 1976). The stronger the nonlinearity, the larger the bias. Second, although long-span samples are now available for many financial variables, making asymptotic properties of econometric estimators more relevant, full

data sets are not always employed in estimation because of possible structural changes in long-span data. When short-span samples are used in estimation, finite-sample distributions can be far from the asymptotic theory. Third, in dynamic models that are used for pricing claims that are contingent on persistent state variables, the MLE of the system parameters may sustain substantial finite-sample bias even in very large samples; and when biased estimated parameters are inserted into the pricing formulas, the bias can be amplified in the resulting estimates of the contingent-claim price. Phillips and Yu (2005) reported evidence of significant bias in the MLE of short-term interest rate models. The present paper shows that bias in the MLE of volatility models can also be substantial, especially in worst-case scenarios where there is persistence and nonlinearity.

Some past studies in the literature have addressed the finite-sample properties of estimators of contingent-claim prices. Boyle and Ananthanarayanan (1977) examined the exact finite-sample distribution of the estimated Black-Scholes option price evaluated at an unbiased estimator of the true variance and showed the resultant estimator to be biased. To remove the bias, Butler and Schachter (1986) proposed an estimator based on Taylor series expansions. Knight and Satchell (1997) showed that the estimator of Butler and Schachter is only unbiased for at-the-money options. When ML is used to estimate one-factor models for short-term interest rates, Ball and Torous (1996) and Chapman and Pearson (2000) provided evidence of large finite-sample biases in the mean reversion parameter. Phillips and Yu (2005) showed that this bias translates into bond pricing and bond option pricing and the pricing biases are economically too significant to ignore. To reduce these biases, Phillips and Yu (2005) proposed a new jackknife procedure. While the method proposed by Butler and Schachter (1986) is fundamentally different from that of Phillips and Yu (2005), they share a common limiting property: relative to ML, both these methods trade off the gain that may be achieved in bias reduction with a loss that arises through increased variance.

The present paper introduces a new methodology of estimating contingent-claim prices that can achieve bias reduction as well as variance reduction, thereby offering overall gains in mean square estimation error for contingent-claim pricing. Instead of inserting a bias-corrected ML estimator into the pricing formulas, the approach involves the direct estimation of contingent-claim prices that is complete with an in-built correction for bias. The proposed method is simulation-based and involves multiple stages. In a preliminary stage, the bias in the price estimator is calibrated via simulation and at the next stage a procedure that accounts for this bias is implemented.

Simulation-based methods have been successfully used in past work to estimate parameters in various financial time-series models. For example, they have been employed in the context of continuous-time models to address issues of discretization bias (e.g., Duffie and Singleton 1993; Monfort 1996; Dai and Singleton 2000) and in the context of discrete-time stochastic volatility models

to deal with intractable likelihoods (e.g., Monfardini 1998; Andersen, Chung, and Lund 1999). The methods have also been utilized to correct finite-sample bias in time-series models (e.g., MacKinnon and Smith 1998; Gourieroux, Renault, and Touzi 2000) and in dynamic panel models (e.g., Gourieroux, Phillips, and Yu 2007). The present work is, to the best of our knowledge, the first implementation of such methods in contingent-claim pricing.

Simulation-based methods have several favorable attributes in the estimation of contingent-claim prices. The first is that they do not require explicit analytic evaluation of the bias function since this function is implicitly calculated by simulation. This advantage is significant as most asset pricing models do not yield analytic expressions for the bias function. Simulation-based methods are therefore applicable in a broad range of model specifications where analytic methods fail.

Second, the simulation approach described here can be used in connection with many different estimation methods, including exact ML when it is available, and various approximate ML techniques. In recent years, building upon the pioneering work of Aït-Sahalia (1999, 2002), an extensive literature has emerged that develops and applies closed-form ML methods to estimate model parameters in various setups. For example, Aït-Sahalia (1999) and Egorov, Li, and Xu (2003) estimated short-term interest rate models. Aït-Sahalia (2007) generalized the technique to multivariate diffusions. Aït-Sahalia and Kimmel (2006); Egorov, Li, and Ng (2008); and Thompson (2008) estimated term structure models. Aït-Sahalia and Kimmel (2007) estimated stochastic volatility models. There are two nice features about the closed-form ML method: (1) the estimator can approximate the exact MLE highly accurately; and (2) being based on a closed analytic form, it is computationally efficient. When the simulation-based method is used in connection with the exact or the closed-form MLE, the resultant estimator is asymptotically equivalent, thereby sharing all the asymptotic properties of the initial MLE, and standard tools of statistical inference are applicable. In this sense, the estimator may be regarded as an extension of the closed-form MLE.

Third, the present methods can deal with both the estimation bias and the discretization bias that arises when nonlinear stochastic differential equations are estimated. Since nonlinear stochastic differential equations typically do not have closed-form likelihood expressions, exact ML estimation presents many challenges. While it is straightforward to estimate a discretized model, discretization bias is inevitably introduced in practice. Simulations permit the sampling interval to be chosen arbitrarily small, thereby providing an important control on the size of the discretization bias. Fourth, simulation-based methods have the advantage of flexibility and can be readily applied in any practical contingent-claim pricing situation.

One drawback of simulation-based methods is that they are inevitably computationally intensive. But numerical methods are now an important aspect of most empirical procedures in finance and ongoing advances in
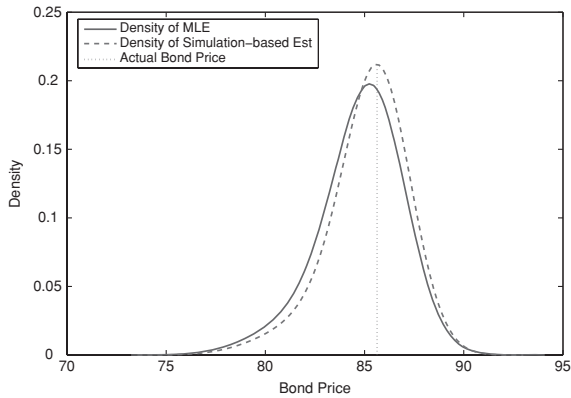
Figure 1
Distributions of simulation-based and ML estimates of bond price for highly persistent data
To obtain the distributions of simulation-based and ML estimates of bond price, we simulate 5000 data sets from the Vasicek model $dS(t) = \kappa(\mu - S(t))dt + \sigma d B(t)$, each with 7500 daily observations (30 years of daily interest rates), and then estimate the price of a three-year discount bond. We choose $\kappa = 0.018$ as a highly dependent case. The graphs show the kernel density of simulation-based and ML estimates of bond price. The solid line is for the MLE; the dashed line is for the simulation-based estimates; the dotted line is the true value.

computing technology continue to make numerically intensive computations less burdensome in practical applications. Moreover, the computational efficiency of the closed-form MLE makes it an ideal initial estimator for our simulation-based methods. Another characteristic of simulation-based methods is that they lack exact reproducibility unless common seeds and random number generators are used. This is because the number of simulation paths is inevitably finite in practical applications.

Our findings here indicate that simulation-based methods provide substantial improvements in pricing contingent claims over ML in the case where ML has a substantial bias. To illustrate, Figure 1 compares the distributions of estimates of the price of a discount bond obtained from 30-year daily data by using the MLE and a bias-corrected simulation method, both in the context of a highly persistent Vasicek model. The actual bond price in this case is $85.63. As is apparent in the figure, the simulation-based estimates are much better centered on the true bond price and achieve bias reduction. In addition, the bias reduction comes with a reduction in variance. In fact, the gain in the percentage bias achieved by the simulation-based method is 64.8% and the gain in standard error is 14.59%. However, when the bias in the MLE is not substantial, ML may well provide the best estimator in finite samples. In this event, simulation-based estimators typically do not provide any improvement over ML. Figure 2 compares the distributions of estimates of the price of a discount bond obtained from 30-year daily data by using the MLE and a bias-corrected simulation method, both in the context of a less persistent Vasicek model. In this case, the
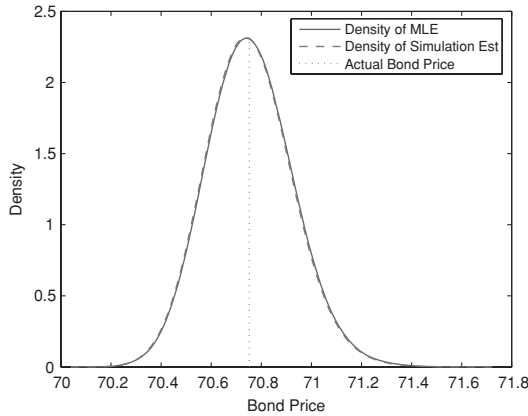
Figure 2
Distributions of simulation-based and ML estimates of bond price for less persistent data
To obtain the distributions of simulation-based and ML estimates of bond price, we simulate 5000 data sets from the Vasicek model $dS(t) = \kappa(\mu - S(t))dt + \sigma dB(t)$, each with 7500 daily observations (30 years of daily interest rates), and then estimate the price of a three-year discount bond. We choose $\kappa = 5$ as a less dependent case. The graphs show the kernel density of simulation-based and ML estimates of bond price. The solid line is for the MLE; the dashed line is for the simulation-based estimates; the dotted line is the true value.

bias in ML is negligible and the two densities are almost identical. More details of this implementation and comparison are provided in Section 2.

While simulation methods can offer improvements over ML when the latter suffers finite-sample problems, ML continues to play an important role for several reasons. First, in many empirically relevant situations, ML does have good finite-sample properties. Second, even in cases where ML may have inferior finite-sample performance, it can still provide a useful first-stage method on which to base the simulation-based methods, as it does here. Third, the good asymptotic behavior of ML will be inherited by suitably designed simulation-based methods that rely on ML.

The paper is organized as follows. Section 1 reviews some existing methods and motivates and introduces our simulation-based methods. Using simulated data, Section 2 explains how the simulation-based methods can be implemented in relation to ML estimation of call options prices in the context of the Black-Scholes model and of bond prices in the context of the Vasicek model. The performance of these simulation-based estimates is compared with that of ML. Section 3 shows how the simulation-based methods can be used to address simultaneously the estimation bias in pricing and the discretization bias. Section 4 examines the practical effects of simulation-based methods in an empirical application with monthly zero-coupon bond data. Section 5 concludes and outlines some further applications and implications of the approach.

1. Estimation Methods for Contingent-Claim Prices

1.1 Maximum likelihood and indirect inference

Let $S(t)$ denote the price of an underlying asset whose dynamics are captured by the following stochastic differential equation:

$$dS(t) = \mu(S(t), t; \theta)dt + \sigma(S(t), t; \theta)dB(t), \qquad (1)$$

where $B(t)$ is a standard Brownian motion, $\sigma(S(t), t; \theta)$ is some specified diffusion function, $\mu(S(t), t; \theta)$ is a given drift function, and θ is an unknown parameter or a vector of unknown parameters. This class of parametric model has been widely used to characterize the temporal dynamics of financial variables, including stock prices, interest rates, and exchange rates.

Although we use a continuous-time model here for $S(t)$, the proposed simulation-based methods will apply more generally to other time-series-generating models for $S(t)$. Unless specified otherwise, the market price of risk is assumed to be zero in this paper. Consequently, the physical measure is identical to the risk-neutral measure.

Suppose a sequence of time-series observations $\mathbf{S} = (S_h, S_{2h}, \ldots, S_{nh})$ taken with a sampling interval h is available over a time period $[0, T(= nh)]$ and we wish to price a financial asset whose payoff is contingent upon the value of $S(t)$. When there is no confusion, we write these observations as $\{S_t\}_{t=1}^n$. Using the no-arbitrage argument, one can derive the price of the contingent claim. Denote by $P(\theta)$ the price of this contingent claim. In general, P may also depend on other parameters that occur in the setting and such dependencies can be accounted for in our approach. But for convenience and exposition, we write P as a function solely of θ.

A common strategy for estimating $P(\theta)$ is to first estimate the parameter vector from the underlying model (such as Equation (1)) based on the data $\mathbf{S}$, leading to the estimate $\hat{\theta}$, and then proceed to insert $\hat{\theta}$ in the pricing function P, giving $\hat{P} = P(\hat{\theta})$.

It has been argued that one should use ML to estimate θ whenever ML is feasible (see Aït-Sahalia 2002 and Durham and Gallant 2002). Since the model (1) has the Markov property, one can write the log-likelihood function as

$$\ell(\theta) = \sum_{t=2}^n \ln f(S_t | S_{t-1}; \theta), \qquad (2)$$

where $f(S_t | S_{t-1})$ denotes the conditional density function of S_t given S_{t-1}. Maximizing the log-likelihood function with respect to θ leads to the MLE $\hat{\theta}_n^{ML}$, which is consistent, asymptotically normal, and asymptotically efficient under usual regularity conditions for stationary dynamic models. In such

circumstances, the limit distribution of $\hat{\theta}_n^{ML}$ is given by

$$\sqrt{n}\left(\hat{\theta}_T^{ML} - \theta\right) \xrightarrow{d} N(0, I^{-1}(\theta)), \tag{3}$$

where $I(\theta)$ is the limiting information matrix, and the MLE is considered optimal in the Hajék-LeCam sense, achieving the Cramér-Rao bound and having the highest possible estimation precision in the limit when $n \to \infty$.

By virtue of the principle of invariance, the MLE of $P(\theta)$ is obtained simply by replacing θ in $P(\theta)$ with $\hat{\theta}_n^{ML}$, leading to $\hat{P}_n^{ML} = P(\hat{\theta}_n^{ML})$.[1] By standard delta method arguments, the following asymptotic behavior for $\hat{P}_n^{ML}$ holds:

$$\sqrt{n}\left(\hat{P}_n^{ML} - P\right) \xrightarrow{d} N(0, V_P), \tag{4}$$

where

$$V_P = \frac{\partial P}{\partial \theta'} I^{-1}(\theta) \frac{\partial P}{\partial \theta}. \tag{5}$$

Since the estimator $\hat{P}_n^{ML}$ is the MLE, it has the highest possible precision when $n \to \infty$, and in consequence, this plug-in estimator has been argued to be the preferred approach (see, for example, Lo 1986).

There are at least two problems with this use of the exact ML approach. First, to calculate the exact MLE, one needs a closed-form expression for $\ln f(S_t|S_{t-1}; \theta)$, which is available only in rare cases. Recent years have witnessed a growing interest in approximating $\ln f(S_t|S_{t-1}; \theta)$ with closed-form approximations. Important contributions include Aït-Sahalia (1999, 2002, 2007) and Aït-Sahalia and Yu (2006). Section 3 describes how to deal with this difficulty in the present context. Second, while $\hat{P}_n^{ML}$ has the highest possible precision asymptotically, it does not necessarily perform the best in finite samples, where it may suffer substantial bias. For example, when the time-series behavior in S_t is highly persistent and $\mu(S_t; \theta)$ is an affine function in S_t, (e.g., $\mu(S_t; \theta) = \kappa(\mu - S_t)$), the exact MLE of κ is substantially upward biased even in large samples (Phillips and Yu 2005). This upward bias in $\widehat{\kappa}_n^{ML}$ translates to bias in $\hat{P}_n^{ML}$ and may be large enough to be of economic significance in practice.

To appreciate the circumstances where $\hat{P}_n^{ML}$ suffers bias, expand $P(\hat{\theta}_n^{ML})$ around θ (assuming $P(\theta)$ to be twice differentiable and θ to be scalar) and, taking expectations, we have the approximate expression

$$E\left(\hat{P}_n^{ML}\right) \approx P(\theta) + E\left(\hat{\theta}_n^{ML} - \theta\right) \frac{\partial P(\theta)}{\partial \theta} + \frac{1}{2} \text{Var}\left(\hat{\theta}_n^{ML}\right) \frac{\partial^2 P(\theta)}{\partial \theta^2}. \tag{6}$$

Equation (6) indicates three situations where $\hat{P}_n^{ML}$ will incur substantial bias: first, when $\hat{\theta}_n^{ML}$ is itself strongly biased; second, when $P(\theta)$ is highly nonlinear

---

[1] While we assume in this paper that P is not a function of the underlying asset price, this assumption may be relaxed and, conditional on the underlying asset price, the principle of invariance still applies.

and $\partial P/\partial \theta$ is large; and third, when $\mathrm{Var}(\hat{\theta}_n^{ML})$ is large, which is typically the case in small-sample situations.

To illustrate the possible finite-sample problems that can arise with MLE, we consider three examples here. These involve worst-case scenarios to illustrate the difficulties. A wider set of examples is given in Section 2 for some of which MLE performs very well. In the first example, we estimate the price of a very deep out-of-the-money option in the context of the Black-Scholes model. In the second example, we estimate a bond price and a bond option price in the context of the Vasicek model. The third example looks at the option price in the context of the stochastic volatility model of Hull and White (1987). Some further details of the examples are given in Section 2.

In the first example, let $S(t)$ be the price of an underlying stock at time t, which is assumed to follow the geometric Brownian motion process (Black and Scholes 1973):

$$dS(t) = \mu S(t)dt + \sigma S(t)dB(t), \tag{7}$$

and let $\{S_t\}_{t=0}^n$ be a sample of equispaced time-series observations on $S(t)$ with sampling interval h and $T = nh$. In the Black-Scholes option pricing formula, the only unknown quantity is σ^2. Since $\hat{\sigma}_n^{2,ML} \equiv \frac{1}{T} \sum_{t=0}^{n-1} (\ln \frac{S_{t+1}}{S_t} - \frac{1}{n} \sum_{t=0}^{n-1} \ln \frac{S_{t+1}}{S_t})^2$ is the MLE of σ^2, $\hat{P}_n^{ML} = P(\hat{\sigma}_n^{2,ML})$ is the MLE of P, an estimator advocated in Lo (1986). Moreover, Lo showed that

$$\sqrt{n}(\hat{P}_T^{ML} - P) \xrightarrow{d} N\left(0, \frac{\tau}{2} S^2 \sigma^2 \phi^2(d_1)\right), \tag{8}$$

where τ is the time to maturity and ϕ is the density of the standard normal distribution. We use 250 (simulated) daily stock returns (i.e., $h = 1/250$) to obtain the ML estimates of σ^2 and the price of a deep out-of-the-money European call option that matures in one week. The experiment is replicated 5000 times to obtain the mean, the percentage bias, and the root mean square error (RMSE). Table 1 reports on the results. It can be seen that while the MLE of σ^2 has little bias (-0.48%), the bias in $\hat{P}_n^{ML}$ is substantial (19.6%) and economically significant. Obviously, in this case the bias in $\hat{P}_n^{ML}$ comes almost entirely from the nonlinearity in the function P.

To understand why the bias is so severe in $\hat{P}_n^{ML}$ when there is little bias in $\hat{\sigma}_n^{2,ML}$, we plot in Figure 3 $\partial \ln P(\sigma^2)/\partial \sigma$ as a function of σ^2 for options with different degree of moneyness. It can be seen that the deep out-of-the-money option is highly nonlinear while the other two options are nearly linear. As a result, the second and third terms on the right-hand side of Equation (6) are negligible when pricing in-the-money and near-money options but non-negligible when pricing a deep-out-of-the-money option. In Section 2, we will provide examples where the bias in $\hat{P}_n^{ML}$ is negligible when the moneyness is in-the-money or near-the-money.

Table 1
Finite-sample properties of MLE of σ^2 and P in the Black-Scholes model

	$\widehat{\sigma}_n^{2,ML}$	$\widehat{P}_n^{ML}$
True	0.40	2.12
Mean	0.3983	2.53
Bias (in %)	−0.4791	19.60
RMSE	0.0365	2.10

The table reports on the true value, the mean, the bias (in percentage), and the RMSE of MLE of σ^2 and P in the Black-Scholes model obtained from simulations. We simulate 5000 data sets from the Black-Scholes model

$$dS(t) = \mu S(t)dt + \sigma S(t)dB(t),$$

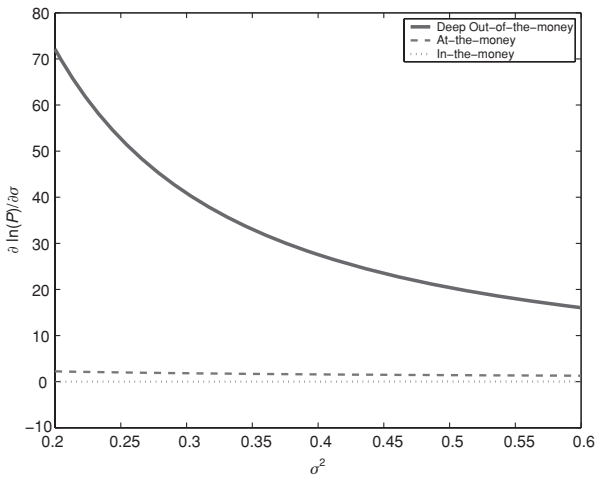each with 250 daily observations.

Figure 3
Derivative of the Black-Scholes price with respect to volatility as a function volatility
The graphs plot $\partial \ln P(\sigma^2)/\partial \sigma$ as a function of σ^2 for options with different degree of moneyness. $P(\sigma^2)$ is the Black-Scholes price defined by $S\Phi(d_1) - Xe^{-r\tau}\Phi(d_2)$, where $d_1 = (\ln(S/X) + (r + 0.5\sigma^2)\tau)/\sigma\sqrt{\tau}$ and $d_2 = d_1 - \sigma\sqrt{\tau}$.

In the second example, the short-term interest rate $S(t)$ is assumed to follow the Ornstein-Uhlenbeck process (Vasicek 1977):

$$dS(t) = \kappa(\mu - S(t))dt + \sigma dB(t), \tag{9}$$

and $\{S_t\}_{t=1}^n$ is a sample of equispaced time-series observations on $S(t)$ over $[0, T(= nh)]$ with sampling interval h. In the Vasicek bond pricing formula, the unknown quantities are κ, μ, and σ^2. It is known that μ and σ^2 can be estimated with little bias by exact ML (Tang and Chen 2007), so we fix these two parameters and let κ be the only unknown parameter in the simulation. We use 7500 (simulated) daily observations (i.e., $h = 1/250$) to obtain the ML estimates of κ, the price of a three-year discount bond (Bond Price or BP

Table 2
Finite-sample properties of MLE of κ, *BP*, and *OP* in the Vasicek model

	$\widehat{\kappa}_n^{ML}$	$\widehat{BP}_n^{ML}$	$\widehat{OP}_n^{ML}$
True	0.018	85.653	4.0485
Mean	0.0598	84.745	3.5726
Bias (in %)	232.26	−1.033	−11.7561
RMSE	0.0922	1.8821	1.0058

The table reports on the true value, the mean, the bias (in percentage), and the RMSE of MLE of κ, *BP*, and *OP* in the Vasicek model obtained from simulations. We simulate 5000 data sets from the Vasicek model

$$dS(t) = \kappa(\mu - S(t))dt + \sigma dB(t),$$

each with 7500 daily observations, and price a three-year discount bond and a two-year European call option written on the discount bond. *BP* is defined in Equation (21) while *OP* is defined in Equation (22).

hereafter), and the price of a two-year European call option on the discount bond (Option Price or OP hereafter). The experiment is replicated 5000 times to obtain the mean, the percentage bias, and the RMSE. Table 2 reports on the results. It can be seen that the bias in κ is substantial and since the true value of $\kappa = 0.018$, this may be interpreted as a manifestation of the near unit root problem. The bias naturally translates into $\widehat{BP}_n^{ML}$ and $\widehat{OP}_n^{ML}$, which is of economic significance (see, for example, Hull 2000).

In the last example, $S(t)$ is a stock price, which is assumed to follow the stochastic volatility (SV) model (Hull and White 1987):

$$dS(t) = \sigma_S S(t)\sigma(t)dB_1(t),$$
$$d\ln\sigma^2(t) = -\kappa\ln\sigma^2(t)dt + \gamma dB_2(t),$$

and $\{S_t\}_{t=1}^n$ is again a sample of equispaced time-series observations on $S(t)$ with sampling interval h. Under certain assumptions, Hull and White (1987) showed that the value of a European call option is the Black-Scholes price integrated over the distribution of the mean volatility. Unfortunately, the option price does not have a closed-form solution. A flexible way for calculating option prices is via Monte Carlo simulations. For example, Hull and White (1987) designed an efficient procedure of carrying out the Monte Carlo simulation to calculate a European call option. In general, the price depends on κ, σ_S, and γ. For the SV model, it is well known that the likelihood function has no closed-form expression (Durham and Gallant 2002; Kim, Shephard, and Chib 1998).[2] Several simulation-based ML methods have been proposed in recent years. In this paper, a discretized version of the SV model is estimated by the ML method of Skaug and Yu (2007). We use 500 (simulated) daily

---

[2] When only price data are used to estimate the SV model, the latent volatility process has to be integrated out from the joint density of prices and volatility, making the evaluation of likelihood numerically demanding. However, when the latent volatility is obtained from option prices, Aït-Sahalia and Kimmel (2007) showed that an approximate ML is feasible. In this case, if volatility is highly persistent, the same finite-sample problem in ML can be expected to occur as for the persistent Vasicek model.

Table 3
Finite-sample properties of MLE of σ_S, κ, γ, and P in the lognormal SV model

	$\widehat{\sigma}_{S,n}^{ML}$	$\widehat{\kappa}_n^{ML}$	$\widehat{\gamma}_n^{ML}$	$\widehat{P}_n^{ML}$
True	5.521	5.000	2.372	0.105
Mean	5.536	9.839	2.573	0.091
Bias (in %)	0.26	96.79	8.47	−13.08
RMSE	0.328	13.054	0.884	0.034

The table reports on the true value, the mean, the bias (in percentage), and the RMSE of MLE of σ_S, κ, γ, and P in the lognormal stochastic volatility model obtained from simulations. We simulate 500 data sets from the lognormal stochastic volatility model of Hull and White (1987),

$$dS(t) = \sigma_S S(t)\sigma(t)dB_1(t),$$
$$d\ln \sigma^2(t) = -\kappa \ln \sigma^2(t)dt + \gamma dB_2(t),$$

each with 500 daily observations.

observations ($h = 1/250$) to obtain the ML estimates of σ_S, κ, γ, and the price of an out-of-the-money European call option.[3] The option prices are calculated based on 1000 simulated paths. The experiment is replicated 500 times to obtain the mean, the percentage bias, and the RMSE. Table 3 reports on the results. As in the Vasicek model, the bias in κ is substantial and is again a manifestation of the near-unit root problem.[4] The bias naturally translates into $\widehat{P}_n^{ML}$ and is economically very significant. To the best of our knowledge, this bias in estimating κ and in pricing seems not to have been noticed in the context of SV models.

All three examples clearly point to a need to correct the bias in ML estimation of contingent-claim prices. The first example suggests that insertion of a bias-corrected parameter estimate into the contingent-claim price does not necessarily work well. The present paper seeks to address this problem by the use of the indirect inference procedure applied *directly* to the contingent-claim price.

Indirect inference is a simulation-based method developed by Smith (1993) for estimating models where the likelihood is difficult to construct analytically but where the model may be readily simulated.[5] It is closely related to the simulated GMM method of Duffie and Singleton (1993) and the efficient method of moments (EMM) technique of Gallant and Tauchen (1996). This method also has the property that it can successfully correct for estimation bias in time-series parameter estimation. The application of indirect inference here proceeds as follows. Let $\widehat{\theta}_T^{ML}$ denote the MLE that is obtained from the actual data and involves some finite-sample estimation bias. For any given parameter choice θ, let $\tilde{\mathbf{S}}^k(\theta) = \{\tilde{S}_1^k, \tilde{S}_2^k, \ldots, \tilde{S}_n^k\}$ be data simulated from the time-series

---

[3] The spot price is $10, the time to maturity for the option contract is 0.5 years, the interest rate is 10%, the strike price is $11.56, and the initial value of σ^2 is 0.02.

[4] Since the sampling interval is very small at the daily frequency, the discretization bias is negligible (Phillips and Yu 2007).

[5] The name indirect inference was coined by Gourieroux, Monfort, and Renault (1993), who further developed the methodology.

model (1), where $k = 1, \ldots, K$ and K is the number of simulated paths. The number of observations in $\tilde{\mathbf{S}}^k(\theta)$ is chosen to be the same as the number of actual observations in $\mathbf{S}$ so that the exact finite-sample properties of $\hat{\theta}_n^{ML}$, including its finite-sample bias, may be calibrated. Let $\tilde{\phi}_n^{ML,k}(\theta)$ denote the MLE of θ obtained in this way from the kth simulated path. By construction, this simulation-based estimate naturally carries any finite-sample estimation bias of the MLE in the given model and for this sample size.

   The idea behind the procedure that leads to bias correction is to choose θ so that the average behavior of $\tilde{\phi}_n^{ML,k}(\theta)$ is matched against the numerical estimate $\hat{\theta}_n^{ML}$ obtained with the observed data. In particular, the indirect inference estimator is defined by

$$\hat{\theta}_{n,K}^{II} = \operatorname{argmin}_{\theta \in \Theta} \left\| \hat{\theta}_n^{ML} - \frac{1}{K} \sum_{k=1}^{K} \tilde{\phi}_n^{ML,k}(\theta) \right\|, \tag{10}$$

where $\| \cdot \|$ is some finite-dimensional distance metric and the region of extremum estimation Θ is a compact set. In the case where K tends to infinity, the law of large numbers, $K^{-1} \sum_{k=1}^{K} \tilde{\phi}_n^{ML,k}(\theta) \to_p E(\tilde{\phi}_n^{ML,k}(\theta))$, applies by virtue of the nature of the simulation and then the indirect inference estimator becomes

$$\hat{\theta}_n^{II} = \operatorname{argmin}_{\theta \in \Theta} \left\| \hat{\theta}_n^{ML} - b_n(\theta) \right\|, \tag{11}$$

where $b_n(\theta) = E(\tilde{\phi}_n^{ML,k}(\theta))$ is called the binding function. When b_n is invertible, the indirect inference estimator may be written directly as

$$\hat{\theta}_n^{II} = b_n^{-1}(\tilde{\theta}_T^{ML}).$$

The procedure essentially builds in a finite-sample bias correction to $\hat{\theta}_n^{ML}$, with the bias being computed directly by simulation. Any bias that occurs in $\hat{\theta}_n^{ML}$ will also be present in the binding function $b_n(\theta)$. Hence, with the bias correction that is built into the inversion functional $\hat{\theta}_n^{II} = b_n^{-1}(\hat{\theta}_n^{ML})$, the estimator $\hat{\theta}_n^{II}$ becomes exactly "b_n-mean-unbiased" for θ. That is, $E(b_n(\hat{\theta}_n^{II})) = b_n(\theta)$. Moreover, in typical cases where $\lim_{n \to \infty} E(\hat{\theta}_n^{ML}) = \theta$ and $\hat{\theta}_n^{ML}$ is asymptotically unbiased, we have $\hat{\theta}_n^{II} \sim \hat{\theta}_n^{ML}$ in the limit as $n \to \infty$. Then, the indirect inference estimator is asymptotically equivalent to the MLE so that $\hat{\theta}_n^{II}$ shares the same good asymptotic properties of $\hat{\theta}_n^{ML}$, while having improved finite-sample performance.

1.2 Direct simulation-based methods of pricing

While the indirect inference estimator of θ, $\hat{\theta}_n^{II}$, may have better finite-sample properties than $\hat{\theta}_n^{ML}$, inserting $\hat{\theta}_n^{II}$ into $P(\theta)$ does not necessarily lead to a better estimator than $\hat{P}_n^{ML}$ due to the nonlinearity in the pricing function. This phenomenon was explicitly addressed in Phillips and Yu (2005), where it was found that jackknifing the quantity of interest clearly performs better than

the method that inserts the jackknife (hence bias-reduced) estimator and the median unbiased (hence bias-corrected) estimator into the pricing formulas. To improve the finite-sample properties of $\hat{P}_n^{ML}$, we propose to apply simulation-based methods directly in the estimation of contingent-claim prices.

We first focus on the case where θ is a scalar. As above, we denote by $\hat{\theta}_n^{ML}$ the MLE of θ that is obtained from the actual data, and write $\hat{P}_n^{ML} = P(\hat{\theta}_n^{ML})$. $\hat{P}_n^{ML}$ involves finite-sample estimation bias due to the nonlinearity of the pricing function P in θ, or the use of the biased estimate $\hat{\theta}_n^{ML}$, or both these effects.

The simulation approach involves the following steps.

(1) Given a value for the contingent-claim price p, compute $P^{-1}(p)$ (call it $\theta(p)$), where $P^{-1}(\cdot)$ is the inverse of the pricing function $P(\theta)$.

(2) Let $\tilde{\mathbf{S}}^k(p) = \{\tilde{S}_1^k, \tilde{S}_2^k, \ldots, \tilde{S}_T^k\}$ be data simulated from the time-series model (1) given $\theta(p)$, where $k = 1, \ldots, K$ with K being the number of simulated paths. As argued above, we choose the number of observations in $\tilde{\mathbf{S}}^k(p)$ to be the same as the number of actual observations in $\mathbf{S}$ for the express purpose of finite-sample bias calibration.

(3) Obtain $\tilde{\phi}_n^{ML,k}(p)$, the MLE of θ, from the kth simulated path, and calculate $\tilde{P}_n^{ML,k}(p) = P(\tilde{\phi}_n^{ML,k}(p))$.

(4) Choose p so that the average behavior of $\tilde{P}_n^{ML,k}(p)$ is matched with $\hat{P}_n^{ML}$ to produce a new bias-corrected estimate.

Whenever bias occurs in $\hat{P}_n^{ML}$ and from whatever source, this bias will also be present in $\tilde{P}_n^{ML,k}(p)$ for the same reasons. Hence, the procedure builds in a finite-sample bias correction directly to correct $\hat{P}_n^{ML}$. The resultant estimator is different from simply inserting a simulation-based estimator of θ into the pricing functional P, because this approach considers the quantity of interest directly.

We propose using two quantities to represent the average behavior of $P(\tilde{\phi}_n^{ML,k}(p))$ as the binding function. The first one is the mean, which corresponds to the indirect inference estimation approach of Smith (1993) and Gourieroux, Monfort, and Renault (1993), while the second is the median, corresponding to the median unbiased estimation approach of Andrews (1993). Of course, the median is more robust to outliers than the mean. Hence, when the distribution of $\hat{P}_n^{ML}$ is highly skewed, it may be preferable to use the median in this approach. In general, however, the binding function cannot be computed analytically in either case and simulations are needed to calculate the binding functions.

If the mean is chosen to be the binding function, the simulation-based estimator is defined as

$$\hat{P}_{n,K}^{SM,1} = \text{argmin}_p \left\| \hat{P}_n^{ML} - \frac{1}{K} \sum_{k=1}^{K} P(\tilde{\phi}_n^{ML,k}(p)) \right\|. \tag{12}$$

In the case where K tends to infinity, this simulation-based estimator becomes

$$\hat{P}_n^{SM,1} = \text{argmin}_p \left\| \hat{P}_n^{ML} - b_{n,1}(p) \right\|, \tag{13}$$

where the binding function $b_{n,1}(p)$ is $E(P(\tilde{\phi}_n^{ML,k}(p)))$. If $b_{n,1}(p)$ is invertible, we then have

$$\hat{P}_n^{SM,1} = b_{n,1}^{-1}(\hat{P}_n^{ML}). \tag{14}$$

If the median is chosen to be the binding function, the simulation estimator is defined as

$$\hat{P}_{n,K}^{SM,2} = \text{argmin}_p \left\| \hat{P}_n^{ML} - \hat{\rho}_{0.5} P(\tilde{\phi}_n^{ML,k}(p)) \right\|, \tag{15}$$

where $\hat{\rho}_\tau$ is the τth sample quantile obtained from $\{P(\tilde{\phi}_n^{ML,1}(p)), \ldots, P(\tilde{\phi}_n^{ML,K}(p))\}$.[6] In the case where K tends to infinity, this simulation-based estimator becomes

$$\hat{P}_n^{SM,2} = \text{argmin}_p \left\| \hat{P}_n^{ML} - b_{n,2}(p) \right\|, \tag{16}$$

where the binding function $b_{n,2}(p)$ is $\rho_{0.5}(P(\tilde{\phi}_n^{ML,k}(p)))$. If $b_{n,2}(p)$ is invertible, we have

$$\hat{P}_n^{SM,2} = b_{n,2}^{-1}(\hat{P}_n^{ML}). \tag{17}$$

Equation (14) implies that $\hat{P}_n^{SM,1}$ is exactly "b_n-mean-unbiased" for θ in the sense that $E(b_{n,1}(\hat{P}_n^{SM,1})) = b_{n,1}(P)$. Similarly, from Equation (17) it can be shown that $\hat{P}_n^{SM,2}$ is exactly "b_n-median-unbiased" for θ in the sense that $\rho_{0.5}(b_{n,2}(\hat{P}_n^{SM,2})) = b_{n,2}(P)$. If $b_{n,1}(P)$ is linear in P, exact "b_n-mean-unbiasedness" implies exact mean unbiasedness, i.e., $E(\hat{P}_n^{SM,1}) = P$. If $b_{n,2}(P)$ is strictly monotonic in P, exact "b_n-median-unbiasedness" implies exact median unbiasedness, i.e., $\rho_{0.5}(\hat{P}_n^{SM,1}) = P$. Thus, the sufficient condition for ensuring exact mean unbiasedness of $\hat{P}_n^{SM,1}$ is stronger than the sufficient condition for ensuring exact median unbiasedness of $\hat{P}_n^{SM,2}$. When $\lim_{n\to\infty} E(\hat{P}_n^{ML}) = \lim_{n\to\infty} \rho_{0.5}(\hat{P}_n^{ML}) = P$ and the slopes of the functions $b_{n,1}(P)$ and $b_{n,2}(P)$ are unity as $n \to \infty$, the two simulation-based estimators $\hat{P}_n^{SM,1}$ and $\hat{P}_n^{SM,2}$ are asymptotically equivalent to the MLE $\hat{P}_n^{ML}$.

Applying the delta method to Equations (14) and (17), we obtain for $i = 1, 2$

$$\hat{P}_n^{SM,i} = b_{n,i}^{-1}(\hat{P}_n^{ML}) = b_{n,i}^{-1}(b_{n,i}(P) + \hat{P}_n^{ML} - b_{n,i}(P)),$$

and

$$\text{Var}(\hat{P}_n^{SM,i}) \approx \left(\frac{\partial b_{n,i}(P)}{\partial P}\right)^{-2} \text{Var}(\hat{P}_n^{ML}) \approx \left(\frac{\partial b_{n,i}(P)}{\partial P}\right)^{-2} \frac{V_P}{n}, \tag{18}$$

---

[6] A number m is the τth quantile of a random variable X if $\text{Prob}(X \geq m) = 1 - \tau$ and $\text{Prob}(X < m) = \tau$. The sample quantile is the sample counterpart of the quantile.

where V_P is given in Equation (5). The asymptotic approximation (18) suggests that the simulation-based estimators should inherit some of the "efficiency" properties of the ML estimator. In fact, the change in the variance depends largely on $\partial b_{n,i}(P)/\partial P$, the slope of the binding function, as seen above. For $|\partial b_{n,i}(P)/\partial P| > 1$, $\hat{P}_n^{SM,i}$ has a smaller variance than the MLE, and for $|\partial b_{n,i}(P)/\partial P| < 1$, $\hat{P}_n^{SM,i}$ has a larger variance than the MLE.

We now consider the case where θ is an M_θ-dimensional vector. Denote by $\hat{\theta}_n^{ML}$ the MLE of θ, obtained from actual data. An important first step in the simulation-based method is to back out θ from contingent-claim prices. To achieve identification, we have to estimate $M_p \geq M_\theta$ contingent-claim prices p to ensure the existence and uniqueness of the inverse mapping $P^{-1}(p)$. These contingent claims may differ in maturities, strike prices, or other features. If the number of contingent claims M_p exceeds M_θ, the inverse $P^{-1}(p)$ will not generally exist unless the equations $p = P(\theta)$ are fully consistent, although we may compute the least squares solution:

$$\theta_{\min} = \underset{\theta}{\operatorname{argmin}} \|P(\theta) - p\|, \quad \|P(\theta) - p\| = (P(\theta) - p)'(P(\theta) - p).$$

If the dimension M_θ of θ outnumbers the contingent claims M_p, then there is generally insufficient information to recover θ from $p = P(\theta)$ and θ is not identified. We will therefore assume in what follows that $M_\theta = M_p$ and that P is invertible. After the inversion, the same steps are used to obtain the simulation-based estimator of P. Since P is now multidimensional, Equation (18) becomes

$$\begin{aligned} \operatorname{Var}(\hat{P}_n^{SM,i}) &\approx \left(\frac{\partial b_{n,i}(P)}{\partial P'}\right)^{-1} \operatorname{Var}(\hat{P}_n^{ML}) \left(\frac{\partial b_{n,i}(P)}{\partial P}\right)^{-1} \\ &\approx \left(\frac{\partial b_{n,i}(P)}{\partial P'}\right)^{-1} \frac{V_P}{n} \left(\frac{\partial b_{n,i}(P)}{\partial P}\right)^{-1}. \end{aligned} \quad (19)$$

To reduce the computation cost, one can choose a fine grid of discrete points, P, from an extended Euclidean space and obtain the binding function on the grid via simulations. Then standard interpolation and extrapolation methods can be used to approximate the binding functions at any point. In this paper, a linear interpolation and extrapolation method is used.

1.3 Simulation-based methods for cross-sectional data
Unlike the time-series case where the gold standard method of estimation is ML, in the cross-section case no single estimation method is regarded as a gold standard. To apply the simulation-based methods, the models have to be assumed for the theoretical contingent-claim prices and for the relation between the theoretical prices and observed prices.

Suppose τ_i and X_i are, respectively, the time-to-maturity and the strike price of an option, where $i = 1, \ldots, n$. Let $\hat{P}_i(\tau_i, X_i)$ be its observed price and

$P_i(\theta; \tau_i, X_i)$ its model price as determined by the option formula corresponding to model (1). Assume that

$$\hat{P}_i = P_i(\theta; \tau_i, X_i) + \epsilon_i, \; \epsilon_i \sim N(0, \sigma_e^2). \tag{20}$$

Under this specification, we can use OLS to estimate θ from cross-sectional data on option prices $\{\hat{P}_i\}_{i=1}^n$, giving

$$\hat{\theta} = \text{argmin}_\theta \sum (\hat{P}_i - P_i(\theta; \tau_i, X_i))^2.$$

An estimate of σ_e^2 is obtained by

$$\hat{\sigma}_e^2 = \frac{1}{n-1} \sum_{i=1}^n (\hat{P}_i - P_i(\hat{\theta}; \tau_i, X_i))^2.$$

To estimate the price of an option with a new time-to-maturity and a new strike price (say, τ_{n+1} and X_{n+1}, respectively), a commonly used estimator in the literature (see, for example, Bakshi, Cao, and Chen 1997) is

$$\hat{P}_{n+1} = P_{n+1}(\hat{\theta}; \tau_{n+1}, X_{n+1}).$$

To implement the simulation-based method, we suggest the following steps:

(1) Given a value for the option price p, compute the implied parameter at the new time-to-maturity and a new strike price. Define the implied parameter by $P_{n+1}^{-1}(p; \tau_{n+1}, X_{n+1})$.

(2) Let $\tilde{\mathbf{P}}^k(p) = (\tilde{P}_1^k(p), \dots, \tilde{P}_n^k(p))$ be data simulated from (20), given $P_{n+1}^{-1}(p; \tau_{n+1}, X_{n+1})$, $\{\tau_i, X_i\}_{i=1}^n$, $\hat{\sigma}_e^2$, where $k = 1, \dots, K$ with K being the number of simulated paths.

(3) From the kth simulated path, obtain the OLS estimate of θ (call it $\tilde{\theta}^k$) and the estimate of P_{n+1} (call it $\tilde{P}_{n+1}^k$).

(4) Choose p so that the average behavior of $\tilde{P}_{n+1}^k(p)$ is matched with $\hat{P}_{n+1}$ to produce a simulation-based estimate.

2. Illustrations and Monte Carlo Evidence

This section illustrates the bias problem in the estimation of contingent-claim prices in the context of both the Black-Scholes option pricing model and the Vasicek bond and option pricing model. The reason for considering these two specific models in the Monte Carlo study is that they both have closed-form expressions for the conditional densities and we can therefore perform exact ML estimation of P, providing a useful benchmark of comparison. Moreover, the contingent-claim prices have closed-form expressions in these two models. We also discuss situations whereby ML does not suffer from bias problems and simulation-based methods do not offer any improvement over ML.

2.1 Black-Scholes option pricing

As shown in Section 1, in the Black-Scholes model, for deep out-of-the-money options with a short time-to-maturity, $\hat{P}_n^{ML}$ can be substantially biased. It is therefore of particular interest to see how bias reduction strategies work in this case.

First, we define the following notation:

X = Strike price,

τ = Time to maturity,

r = Interest rate,

$\hat{\sigma}_n^{2,ML}$ = MLE of σ^2 defined by $\frac{1}{T}\sum_{t=0}^{n-1}(\ln\frac{S_{t+1}}{S_t} - \frac{1}{n}\sum_{t=0}^{n-1}\ln\frac{S_{t+1}}{S_t})^2$,

s_n^2 = Bias-corrected MLE of σ^2 defined by $\frac{n}{n-1}\hat{\sigma}_T^{2,ML}$,

d_1 = $\frac{1}{\sigma\sqrt{\tau}}(\ln(S/X) + (r + 0.5\sigma^2)\tau)$,

d_2 = $d_1 - \sigma\sqrt{\tau}$,

Φ = Cumulative distribution function of standard normal distribution,

P = Price of a European call option obtained from $S\Phi(d_1) - Xe^{-r\tau}\Phi(d_2)$.

In finite samples, however, $\hat{\sigma}_n^{2,ML}$ is slightly biased, while s_n^2 is unbiased. Inserting s_n^2 into $P(\sigma^2)$ is an alternative estimator of P that has received a great deal of attention (see, for example, Boyle and Ananthanarayanan 1977 and Butler and Schachter 1986). In particular, Boyle and Ananthanarayanan (1977) obtained exact finite-sample moments of $P(s_n^2)$ and showed that $P(s_n^2)$ is a biased estimator of P. They further provided evidence of the small magnitude of the bias for near- and in-the-money options. However, when the option is deep out-of-the-money, the size of the bias becomes large. Based on a Taylor series expansion of the cumulative distribution function of the standard normal distribution and the distribution of the minimum variance unbiased estimator of σ^2, Butler and Schachter (1986) derived an unbiased estimator of P. Knight and Satchell (1997) showed that a uniformly minimum variance unbiased estimator of P exists if and only if the option is at-the-money.

We now compare the performance of some existing methods with the proposed simulation-based methods using simulated data. In particular, a simple simulation study is conducted to compare the performance of $\hat{P}_n^{ML}$, $P(s_n^2)$, $\hat{P}_n^{SM,1}$, and $\hat{P}_n^{SM,2}$. Throughout the simulations, the following parameter values are used:

$S = \$100$

$\tau = 5/250$

$r = 5\%$

$n = 250$

$h = 1/250$.

That is, we use 250 daily stock returns to estimate the price of a European call option that matures in one week and obtain the estimates $\hat{P}_n^{ML}$, $P(s_n^2)$, $\hat{P}_n^{SM,1}$, and $\hat{P}_n^{SM,2}$. The experiment is replicated 5000 times to obtain the

Table 4
Finite-sample properties of $\hat{P}_n^{ML}$, $P(s_n^2)$, $\hat{P}_n^{SM,1}$, and $\hat{P}_n^{SM,2}$ in the Black-Scholes model for an at-the-money option and an in-the-money option

Estimators	At-the-money option price True value $P = 3.5671$				In-the-money option price True value $P = 6.5341$			
	$\hat{P}_n^{ML}$	$P(s_n^2)$	$\hat{P}_n^{SM,1}$	$\hat{P}_n^{SM,2}$	$\hat{P}_n^{ML}$	$P(s_n^2)$	$\hat{P}_n^{SM,1}$	$\hat{P}_n^{SM,2}$
Mean	3.5548	3.5619	3.5671	3.5679	6.5250	6.5308	6.5341	6.5358
Bias (in %)	−0.3440	−0.1442	0.00	0.0249	−0.1396	−.0495	0.00	0.0268
Std err	0.1625	0.1628	0.1631	0.1631	0.1342	0.1345	0.1348	0.1348
RMSE	0.1630	0.1629	0.1631	0.1631	0.1345	0.1346	0.1348	0.1349
Median	3.5539	3.5610	3.5662	3.5671	6.5232	6.5291	6.5323	6.5341

The table reports on the mean, the bias (in percentage), the standard error, the RMSE, and the median of $\hat{P}_n^{ML}$, $P(s_n^2)$, $\hat{P}_n^{SM,1}$, and $\hat{P}_n^{SM,2}$ in the Black-Scholes model obtained from simulations. We simulate 5000 data sets from the Black-Scholes model

$$dS(t) = \mu S(t)dt + \sigma S(t)dB(t),$$

each with 250 daily observations. The strike price X is set to be $0.95 \times S \exp(r\tau)$ (in-the-money) and $S \exp(r\tau)$ (at-the-money), respectively, and $\sigma^2 = 0.4$.

means, standard errors, RMSEs, and medians of all four estimates. For the two simulation-based estimates, we choose the number of simulated paths to be $K = 5000$. It is well known that σ^2 can be accurately estimated from daily data. Hence, we expect little finite-sample bias in $\hat{\sigma}_n^{2,ML}$.

Table 4 shows the results when $\sigma^2 = 0.4$ and $X = 0.95 \times S \exp(r\tau)$ (in-the-money), $X = S \exp(r\tau)$ (at-the-money), respectively. In this case, the actual option prices are \$6.5341 and \$3.5671. Several conclusions can be drawn from the results reported in the table. First, consistent with what has been documented, we found that $\hat{P}_n^{ML}$ has very small percentage bias (−0.3440% and −0.1396%). Moreover, it has the smallest variance among the four estimators. This is not surprising since σ^2 can be accurately estimated and there is no strong nonlinearity in the options. Second, compared with $\hat{P}_n^{ML}$, the use of an unbiased plug-in estimator, $P(s_n^2)$, reduces the percentage bias to −0.1442% and −0.0495%, respectively, but slightly increases the variance. Third, the bias is further reduced by the simulation estimators $\hat{P}_n^{SM,1}$ (0%) and $\hat{P}_n^{SM,2}$ (0.0249% and 0.0268%). Note that $\hat{P}_n^{SM,1}$ is exactly mean-unbiased and $\hat{P}_n^{SM,2}$ is exactly median-unbiased. While both simulation methods offer bias reduction over $\hat{P}_n^{ML}$, they also marginally increase the variance. Finally, all four estimators perform similarly in terms of RMSE. It is not surprising that the two simulation estimators do not improve over ML because there is little bias in $\hat{\sigma}_n^{2,ML}$ and there is little nonlinearity in $P(\sigma^2)$. Hence, by design ML is expected to have good finite-sample properties in this case and the simulation estimators have very similar performance.

To help understand the performance of the two simulation-based methods, we plot the binding functions in Figure 4 for the at-the-money option. Several features are apparent in the figure. First, both binding functions are very close to the 45° line, suggesting that only a small amount of bias correction is needed in
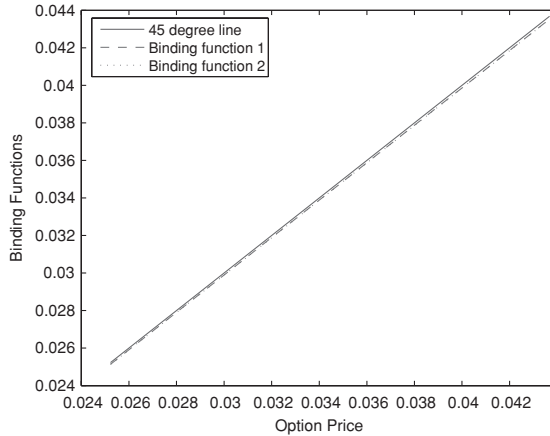
Figure 4
Binding functions of the two simulation-based methods for the at-the-money option
The graphs plot $b_{n,i}(P)$ as a function of p for $i = 1, 2$, respectively, obtained from the Black-Scholes model for an at-the-money option, with 5000 data sets simulated, each with 250 daily observations. The dashed line is for $b_{n,1}(P)$; the dotted line for $b_{n,2}(P)$. The 45° line is plotted for comparison.

the two simulation-based methods in this case. Second, both binding functions are virtually linear, implying that $\hat{P}_n^{SM,1}$ should be exactly mean unbiased and $\hat{P}_n^{SM,2}$ should be exactly median unbiased, a result consistent with the Monte Carlo findings in Table 4. Third, the slopes of the two binding functions are close to but slightly less than 1, suggesting that the variances of the two simulation-based estimators are close to, but slightly larger than, that of $\hat{P}_n^{ML}$. This finding also corroborates the results found in Table 4.

In light of the finding in the literature that standard estimation methods tend to generate large percentage biases for deep out-of-the-money options, we designed an experiment to compare the performance of the four methods when $X = 1.4S \exp(r\tau)$ (i.e., a deep out-of-the-money option) and $\sigma^2 = 0.4$. This is of course a worst-case scenario but may be practically relevant for some stocks. Table 5 reports on the means, standard errors, RMSEs, and medians, each multiplied by 1000, of all four estimates across 5000 replications. The actual call option price (multiplied by 10,000) is $2.12.

Several findings emerge from Table 5. First, consistent with findings in the literature, $\hat{P}_n^{ML}$ has a large percentage bias (19.60%) even though the bias in $\hat{\sigma}_n^{2,ML}$ is very small. Moreover, this estimator no longer has the smallest variance. Second, compared with $\hat{P}_n^{ML}$, instead of reducing the bias, $P(s_n^2)$ increases the percentage bias to 23.51%, so the effect of plugging in an unbiased estimator increases bias. This is a typical example where plugging the bias-corrected ML estimator into the contingent-claim price does not necessarily lead to desirable finite-sample properties. Third and most importantly, the bias is reduced in $\hat{P}_n^{SM,1}$ (1.13%) in terms of the mean and in $\hat{P}_n^{SM,2}$ (0%) in terms of the median. The performance of $\hat{P}_n^{SM,1}$ is particularly encouraging. This estimate not only reduces the bias in terms of the mean, but also decreases variance,

Table 5
Finite-sample properties of $\hat{P}_n^{ML}$, $P(s_n^2)$, $\hat{P}_n^{SM,1}$, and $\hat{P}_n^{SM,2}$ in the Black-Scholes model for a deep out-of-the-money option

| Estimators | Estimation of deep-out-of-the-money option price True value $P = 2.12$ | | | |
	$\hat{P}_n^{ML}$	$P(s_n^2)$	$\hat{P}_n^{SM,1}$	$\hat{P}_n^{SM,2}$
Mean	2.53	2.61	2.14	2.68
Bias (in %)	19.60	23.51	1.13	26.91
Std error	2.05	2.11	1.83	2.16
RMSE	2.10	2.17	1.83	2.24
Median	1.99	2.06	1.64	2.12

The table reports on the mean, the bias (in percentage), the standard error, the RMSE, and the median of $\hat{P}_n^{ML}$, $P(s_n^2)$, $\hat{P}_n^{SM,1}$, and $\hat{P}_n^{SM,2}$ in the Black-Scholes model obtained from simulations. We simulate 5000 data sets from the Black-Scholes model

$$dS(t) = \mu S(t)dt + \sigma S(t)dB(t),$$

each with 250 daily observations. The strike price X is set to be $1.4 \times S \exp(r\tau)$ (i.e., a deep out-of-the-money option) and $\sigma^2 = 0.4$.
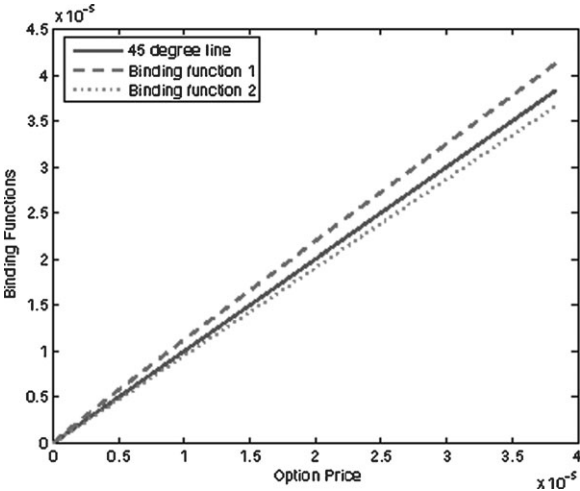

Figure 5
Binding functions of the two simulation-based methods for the deep out-of-the-money option
The graphs plot $b_{n,i}(P)$ as a function of P for $i = 1, 2$, respectively, obtained from the Black-Scholes model for a deep out-of-the-money option, with 5000 data sets simulated, each with 250 daily observations. The dashed line is for $b_{n,1}(P)$; the dotted line for $b_{n,2}(P)$. The 45° line is plotted for comparison.

producing a substantial overall gain in RMSE over $\hat{P}_n^{ML}$ and $P(s_n^2)$. The percentage reductions in RMSE by $\hat{P}_n^{SM,1}$ are 14.75% and 18.58%, respectively, over $\hat{P}_n^{ML}$ and $P(s_n^2)$. Figure 5 plots the two binding functions in this case. The second binding function appears linear and hence monotonic. Not surprisingly, we found evidence of median unbiasedness in $\hat{P}_n^{SM,2}$ in the Monte Carlo study. However, the slope of this binding function is clearly less than 1, leading to an increase in the variance of $\hat{P}_n^{SM,2}$ over $\hat{P}_n^{ML}$. The first binding function has a slope noticeably greater than 1. That explains the finding that $\hat{P}_n^{SM,1}$ reduces the variance.

2.2 Vasicek bond and bond option pricing

Vasicek (1977) introduced a simple term structure model of interest rates where the short-term interest rate is assumed to follow the Ornstein-Uhlenbeck process (9). In process (9), $S(t)$, the short-term interest rate, mean-reverts toward the unconditional mean μ and κ measures the speed of the reversion. Vasicek (1977) derived the expression for bond prices and Jamshidian (1989) gave the corresponding formula for bond option prices. Define $BP(t, s)$ as the price at time t of a discount bond that pays off \$100 at time s and $OP(t, \tau; s, K)$ as the value at time t of a call option on a discount bond of maturity data s and of principal $L = 100$, with exercise (or strike) price K and expiration date τ $(s > \tau > t)$. Vasicek (1977) showed that[7]

$$BP(t, s) = 100 \times A(t, s)e^{-B(t,s)r}, \tag{21}$$

and Jamshidian (1989) showed that

$$OP(t, s; \tau, K) = L \times BP(t, s)\Phi(h) - K \times BP(t, \tau)\Phi(h - \sigma_p), \tag{22}$$

where

$$B(t, s) = \frac{1 - e^{-\kappa(s-t)}}{\kappa},$$

$$A(t, s) = \exp\left(\frac{(B(t, s) - s + t)(\kappa^2\mu - \sigma^2/2)}{\kappa^2} - \frac{\sigma^2 B^2(t, s)}{4\kappa}\right),$$

$$\sigma_p = \frac{\sigma}{\kappa}\left(1 - e^{-\kappa(\tau-s)}\right)\sqrt{\frac{1 - \exp(-2\kappa t)}{2\kappa}},$$

$$h = \frac{1}{\sigma_p}\frac{L \times P(t, \tau)}{P(t, s) \times K} + \frac{\sigma_p}{2},$$

r the initial interest rate, $\Phi(\cdot)$ the cumulative distribution function of standard normal. Obviously, both the bond price BP and the option price OP are functions of $\theta = (\kappa, \mu, \sigma^2)'$, the parameters in Equation (9).

When a discrete sample of the short-term interest rate is available, the exact ML can be used to estimate the parameters. In particular, the conditional density is given by Phillips (1972):

$$S_{t+1}|S_t \sim N(\mu(1 - e^{-\kappa h}) + e^{-\kappa h} S_t, \sigma^2(1 - e^{-2\kappa h})/(2\kappa)). \tag{23}$$

Since κ usually takes a small, positive value and h is often small (1/12, 1/52, and 1/250 for monthly, weekly, and daily data, respectively), $e^{-\kappa h}$ can be well approximated by $1 - \kappa h$. Hence, the Vasicek model is equivalent to a local-to-unity discrete-time autoregressive model, with κ being the local-to-unity parameter, which is well known to be difficult to estimate. Indeed, the ML

---

[7] The formulas given in Equations (21) and (22) are based on the assumption that the market price of risk is zero and have to be adjusted when the market price of risk is different from zero.

estimator of κ is severely upward biased and such a bias translates to bond pricing (see, for example, Phillips and Yu 2005). For example, using 600 monthly observations to estimate the bond and bond option prices, Phillips and Yu (2005, Table 5) found that ML underestimates the price of a three-year discount bond by 1.84% and the price of a one-year bond option by 36.2%. These biases are large and economically important.

We now compare the performance of the MLEs of *BP* and *OP* with the proposed simulation-based estimators in a Monte Carlo study. It is known that μ and σ^2 can be estimated with little bias by exact ML, so we fix these two parameters and let κ be the only unknown parameter in the simulation. For the Vasicek model with known μ, Yu (2008) derived the following formula to approximate the bias of the MLE of κ:[8]

$$E(\hat{\kappa}) - \kappa = \frac{2}{T}\left(1 + \frac{1}{n}\right) + \frac{1}{2T}(e^{2kh} - 1) + o(T^{-1}). \tag{24}$$

Throughout the simulations, the following parameter values are used:

$r = 5\%$
$n = 7500$
$t = 0$
$\tau = 2$
$s = 3$
$h = 1/250$
$\mu = 0.12$
$L = 100$
$K = 105 \times \exp(-(s - t)r)$
$\sigma = 0.01.$

That is, we use 7500 simulated daily observations (30 years of daily interest rates) to estimate the price of a three-year discount bond and that of a two-year European call option written on the bond. Based on 7500 daily interest rates, $\widehat{BP}_n^{ML}(= BP(\hat{\kappa}_n^{ML}))$, $\widehat{BP}_n^{SM,1}$, $\widehat{BP}_n^{SM,2}$, $\widehat{OP}_n^{ML}(= OP(\hat{\kappa}_n^{ML}))$, $\widehat{OP}_n^{SM,1}$, and $\widehat{OP}_n^{SM,2}$ are all obtained. In addition, we consider two alternative estimation methods. The first alternative method is the jackknife method of Phillips and Yu (2005), applying to the quantity of interest directly and using two subsamples. Denote these estimators by $\widehat{BP}_n^{J}$ and $\widehat{OP}_n^{J}$. The second alternative estimator is the plug-in method, which inserts the bias-corrected ML estimator into the pricing formulas. That is, we first obtain the bias-corrected estimator of κ, based on Equation (24), and then plug this bias-corrected estimator into the pricing formulas. Denote these estimators by $\widehat{BP}_n^{P}$ and $\widehat{OP}_n^{P}$. We replicate the experiment 5000 times to obtain the means, standard errors, RMSEs, and

---

[8] Tang and Chen (2007) derived a similar formula to approximate the bias of the MLE of κ when μ is unknown in the Vasicek model and in the square-root model.

Table 6
Finite-sample properties of $\widehat{BP}_n^{ML}$, $\widehat{BP}_n^{P}$, $\widehat{BP}_n^{J}$, $\widehat{BP}_n^{SM,1}$, and $\widehat{BP}_n^{SM,2}$ in the Vasicek model for highly persistent cases

Estimation of price of a discount bond

| | True value BP =85.630 True value $\kappa = 0.018$ | | | | | True value BP =85.0301 True value $\kappa = 0.0416$ | | | | |
Method	$\widehat{BP}_n^{ML}$	$\widehat{BP}_n^{P}$	$\widehat{BP}_n^{J}$	$\widehat{BP}_n^{SM,1}$	$\widehat{BP}_n^{SM,2}$	$\widehat{BP}_n^{ML}$	$\widehat{BP}_n^{P}$	$\widehat{BP}_n^{J}$	$\widehat{BP}_n^{SM,1}$	$\widehat{BP}_n^{SM,2}$
Mean	84.75	85.28	85.09	85.32	85.08	84.03	84.89	84.49	84.75	84.48
Bias (%)	−1.03	−0.413	−0.632	−0.363	−0.640	−1.175	−0.454	−0.634	−0.335	−0.653
Std err	1.661	1.694	1.827	1.450	1.470	1.842	1.859	2.059	1.737	1.732
RMSE	1.882	1.732	1.905	1.483	1.577	2.095	1.913	2.129	1.761	1.819
Median	85.16	85.92	85.43	85.89	85.63	84.44	85.60	84.81	85.39	85.03

The table reports on the true value, the mean, the bias (in percentage), the standard error, the RMSE, and the median of $\widehat{BP}_n^{ML}$, $\widehat{BP}_n^{P}$, $\widehat{BP}_n^{J}$, $\widehat{BP}_n^{SM,1}$, and $\widehat{BP}_n^{SM,2}$ in the Vasicek model for two highly persistent cases obtained from simulations. We simulate 5000 data sets from the Vasicek model

$$dS(t) = \kappa(\mu - S(t))dt + \sigma dB(t),$$

each with 7500 daily observations and price a three-year discount bond. BP is defined in Equation (21).

medians of all five estimates. For the two simulation-based estimates, we chose K=5000 simulated paths.

Tables 6 and 7 show the results, when the true values of κ are 0.018 and 0.0416, for the bond price and the bond option price, respectively. These values are obtained from the empirical work of Phillips and Yu (2005) and Tang and Chen (2007) and imply that the data are very persistent. Indeed, $\kappa = 0.018$ is the empirical estimate obtained from the U.S. data by Tang and Chen (2007), while $\kappa = 0.0416$ is the empirical estimate obtained from another U.S. data set by Phillips and Yu (2005). The corresponding bond prices are \$85.6298 and \$85.0301, and the corresponding option prices are \$4.0485 and \$3.7079.

Several conclusions can be drawn from Tables 6 and 7. First, consistent with the findings in Phillips and Yu (2005), $\widehat{BP}_n^{ML}$ and $\widehat{OP}_n^{ML}$ are downward biased. In particular, the biases are −1.03% and −1.18% for $\widehat{BP}_n^{ML}$ and −11.76% and −14.08% for $\widehat{OP}_n^{ML}$, when the actual values of κ are 0.018 and 0.0416, respectively. The sizes of these estimation biases are of economic significance. The bias in pricing arises from the substantial bias in the MLE of κ and the nonlinearity in the pricing relations. The substantial bias in the MLE of κ is due to the strong persistence in the data. Moreover, $\widehat{BP}_n^{ML}$ and $\widehat{OP}_n^{ML}$ do not possess the smallest variance. While ML has smaller variance than the jackknife and the unbiased plug-in estimators in all cases, its variance is always larger than the two simulation-based estimators. Second, although the jackknife reduces the bias in the ML estimator in all cases, a substantial portion of the bias still remains. In particular, the biases are −0.632%, and −0.634% for $\widehat{BP}_n^{J}$ and −8.12%, and −98.96% for $\widehat{OP}_n^{J}$, when the actual values of κ are 0.018 and 0.0416, respectively. Also, the bias reduction comes at the cost of the increased variance. Similar quantitative results apply to the bias-corrected plug-in method. Third,

Table 7
Finite-sample properties of $\widehat{OP}_n^{ML}$, $\widehat{OP}_n^{P}$, $\widehat{OP}_n^{J}$, $\widehat{OP}_n^{SM,1}$, and $\widehat{OP}_n^{SM,2}$ in the Vasicek model for highly persistent cases

<div align="center">Estimation of price of a bond option</div>

| Method | True value $OP = 4.0485$ | | | | | True value $OP = 3.7079$ | | | | |
| | True value $\kappa = 0.018$ | | | | | True value $\kappa = 0.0416$ | | | | |
	$\widehat{OP}_n^{ML}$	$\widehat{OP}_n^{P}$	$\widehat{OP}_n^{J}$	$\widehat{OP}_n^{SM,1}$	$\widehat{OP}_n^{SM,2}$	$\widehat{OP}_n^{ML}$	$\widehat{OP}_n^{P}$	$\widehat{OP}_n^{J}$	$\widehat{OP}_n^{SM,1}$	$\widehat{OP}_n^{SM,2}$
Mean	3.57	3.77	3.72	3.87	3.76	3.19	3.42	3.38	3.54	3.43
Bias (%)	−11.76	−6.96	−8.12	−4.50	−7.21	−14.08	−7.90	−8.96	−4.56	−7.61
Std err	0.886	0.888	0.981	0.808	0.790	0.963	0.950	1.086	0.957	0.912
RMSE	1.006	.934	1.034	0.828	0.842	1.095	0.995	1.136	0.971	0.955
Median	3.78	4.04	3.91	4.19	4.05	3.38	3.69	3.54	3.89	3.71

The table reports on the true value, the mean, the bias (in percentage), the standard error, the RMSE, and the median of $\widehat{OP}_n^{ML}$, $\widehat{OP}_n^{P}$, $\widehat{OP}_n^{J}$, $\widehat{OP}_n^{SM,1}$, and $\widehat{OP}_n^{SM,2}$ in the Vasicek model for two highly persistent cases obtained from simulations. We simulate 5000 data sets from the Vasicek model

$$dS(t) = \kappa(\mu - S(t))dt + \sigma dB(t),$$

each with 7500 daily observations and price a two-year European call option written on the three-year discount bond. OP is defined in Equation (22).

in all cases the bias is reduced in the two simulation-based estimators both in terms of the mean as well as the median. This reduction is achieved without sacrificing the RMSE and even the variance. For example, when $\kappa = 0.018$ (0.0416), the standard error of $\widehat{BP}_n^{SM,1}$ is 14.59% (6%) smaller than that of $\widehat{BP}_n^{ML}$, while the standard error of $\widehat{OP}_n^{SM,1}$ is 9.71% (0.6%) smaller than that of $\widehat{OP}_n^{ML}$. As a result, there is a substantial gain in the RMSE. Similar conclusions can be drawn on $\widehat{BP}_n^{SM,2}$ and $\widehat{BP}_n^{SM,2}$. Comparison of the medians with the actual bond prices suggests that $\widehat{BP}_n^{SM,2}$ and $\widehat{OP}_n^{SM,2}$ are median unbiased. Not surprisingly, all the gains in RMSE are substantial. Figure 1 compares the nonparametric densities of $\widehat{BP}_n^{SM,2}$ when the actual true value of $\kappa = 0.018$. As is apparent in the figure, the simulation-based estimates are better centered on the true bond price and the bias reduction is accompanied by a reduction in variance.

To understand why the simulation-based method can reduce the variance of the MLE, Figures 6 and 7 plot the binding functions for the bond price and the option price, respectively. In both cases, the two binding functions involve some interesting nonlinearity. In particular, while the binding functions are more linear in the lower-left corner, they become more nonlinear in the upper-right corner. Note that the upper-right corner corresponds to the range of low values of κ, a region that is empirically more realistic. Since the second binding function is still monotonic, $\widehat{BP}_n^{SM,2}$ and $\widehat{OP}_n^{SM,2}$ remain median unbiased. However, the nonlinearity in the first binding function implies that $\widehat{BP}_n^{SM,1}$ and $\widehat{OP}_n^{SM,1}$ are not mean unbiased. Moreover, in the upper-right corner, the slopes of the two binding functions are larger than unity, explaining why the variance can be
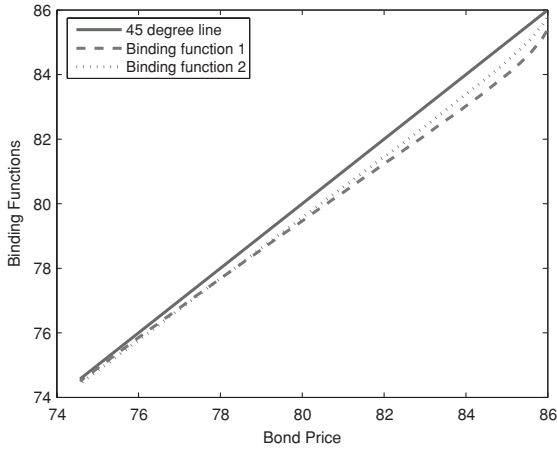
Figure 6
Binding functions of the two simulation-based methods for bond prices under the highly persistent Vasicek model
The graphs plot $b_{n,i}(BP)$ as a function of BP for $i = 1, 2$, respectively, obtained from the Vasicek model for a three-year discount bond, with 5000 data sets simulated, each with 7500 daily observations (30 years of daily interest rates). The dashed line is for $b_{n,1}(BP)$; the dotted line for $b_{n,2}(BP)$. The $45°$ line is plotted for comparison.

Figure 7
Binding functions of the two simulation-based methods for bond option prices under the highly persistent Vasicek model
The graphs plot $b_{n,i}(OP)$ as a function of OP for $i = 1, 2$, respectively, obtained from the Vasicek model for a two-year European option written on a three-year discount bond, with 5000 data sets simulated, each with 7500 daily observations (30 years of daily interest rates). The dashed line is for $b_{n,1}(OP)$; the dotted line for $b_{n,2}(OP)$. The $45°$ line is plotted for comparison.
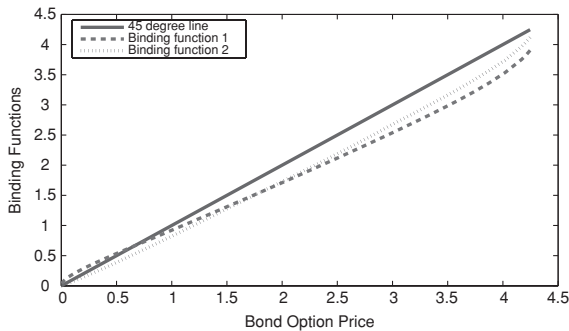
reduced in $\widehat{BP}_n^{SM}$ and $\widehat{OP}_n^{SM}$. All the results are confirmed in the Monte Carlo study.

To appreciate the empirical relevance of the standard error of the bond price estimates, imagine a situation where the interest rate $S(t)$ is 6% and market participants anticipate a change in interest rate in response to three possible movements of the target interest rate (namely, 0, 25, or 50 basis points). Assume that the interest rate has a mean of 6% and a standard error of 25 basis points. Since the three-year BP is related to $S(t)$ by $BP_t = -100 \times \ln S(t)/3$, under

Table 8
Finite-sample properties of $\widehat{BP}_n^{ML}$, $\widehat{BP}_n^{P}$, $\widehat{BP}_n^{J}$, $\widehat{BP}_n^{SM,1}$, and $\widehat{BP}_n^{SM,2}$ in the Vasicek model for intermediate persistent cases

Estimation of price of a discount bond

Method	True value $BP = 75.852$ True value $\kappa = 0.75$					True value $BP = 74.572$ True value $\kappa = 1$				
	$\widehat{BP}_n^{ML}$	$\widehat{BP}_n^{P}$	$\widehat{BP}_n^{J}$	$\widehat{BP}_n^{SM,1}$	$\widehat{BP}_n^{SM,2}$	$\widehat{BP}_n^{ML}$	$\widehat{BP}_n^{P}$	$\widehat{BP}_n^{J}$	$\widehat{BP}_n^{SM,1}$	$\widehat{BP}_n^{SM,2}$
Mean	75.71	76.17	75.81	75.87	75.93	74.50	74.83	74.56	74.58	74.66
Bias (%)	−0.188	0.422	−0.050	0.021	0.104	−0.091	0.346	−0.023	0.014	−0.119
Std err	1.329	1.499	1.408	1.430	1.399	1.048	1.182	1.090	1.105	1.106
RMSE	1.337	1.533	1.408	1.430	1.402	1.050	1.210	1.091	1.105	1.109
Median	75.63	76.11	75.71	75.77	75.85	74.42	74.75	74.45	74.49	74.57

The table reports on the true value, the mean, the bias (in percentage), the standard error, the RMSE, and the median of $\widehat{BP}_n^{ML}$, $\widehat{BP}_n^{P}$, $\widehat{BP}_n^{J}$, $\widehat{BP}_n^{SM,1}$, and $\widehat{BP}_n^{SM,2}$ in the Vasicek model for two intermediate persistent cases obtained from simulations. We simulate 5000 data sets from the Vasicek model

$$dS(t) = \kappa(\mu - S(t))dt + \sigma dB(t),$$

each with 7500 daily observations and price a three-year discount bond. *BP* is defined in Equation (21).

normality the mean of the bond price is then approximately \$93 and the standard error is approximately 1.4. This standard error is of the same order of magnitude as that found using 7500 daily data, as reported in Table 6.

If κ is far away from zero and hence the root is not so near unity, the MLE is expected to have better finite-sample properties. The further the root is from unity, the less bias there will be in the MLE. To examine the performance of ML and the simulation-based methods in this case, we use 7500 daily observations, simulated from the Vasicek model with $\kappa = 0.75$, 1, 5, and 10, respectively, to estimate the price of a three-year discount bond. Based on 7500 daily interest rates, $\widehat{BP}_n^{ML}$, $\widehat{BP}_n^{P}$, $\widehat{BP}_n^{J}$, $\widehat{BP}_n^{SM,1}$, and $\widehat{BP}_n^{SM,2}$ are all obtained. We replicate the experiment 5000 times to obtain the means, standard errors, RMSEs, and medians of all five estimates. For the two simulation-based estimates, we used $K = 5000$ simulated paths.

Table 8 shows the results when the true values of κ are 0.75 and 1. Table 9 shows the results when the true values of κ are 5 and 10. While these values of κ are not empirically realistic for interest rate data in the U.S., they may be relevant for interest rate data in other countries. As expected, $\widehat{BP}_n^{ML}$ is much less downward biased (−0.188%, −0.091%, 0.05%, and 0.02%) when the actual values of κ are 0.75, 1, 5, and 10, respectively. When κ is 0.75 and 1, bias correction methods still manage to reduce the bias despite the small bias of ML. However, there is no gain in RMSE in any of the bias-correction estimators. When κ is 5 and 10, ML is essentially unbiased and the bias correction methods fail to offer any improvement in terms of either bias or RMSE. In these cases, ML should be the method of the choice. Figure 2 compares the nonparametric densities of $\widehat{BP}_n^{SM,1}$ when the actual true value of $\kappa = 5$. As is apparent in the figure, the simulation-based estimates are almost identical to MLE, which is

Table 9
Finite-sample properties of $\widehat{BP}_n^{ML}$, $\widehat{BP}_n^{P}$, $\widehat{BP}_n^{J}$, $\widehat{BP}_n^{SM,1}$, and $\widehat{BP}_n^{SM,2}$ in the Vasicek model for least persistent cases

Estimation of price of a discount bond

| | True value $BP = 70.752$ | | | | | True value $BP = 70.258$ | | | | |
| | True value $\kappa = 5$ | | | | | True value $\kappa = 10$ | | | | |
Method	$\widehat{BP}_n^{ML}$	$\widehat{BP}_n^{P}$	$\widehat{BP}_n^{J}$	$\widehat{BP}_n^{SM,1}$	$\widehat{BP}_n^{SM,2}$	$\widehat{BP}_n^{ML}$	$\widehat{BP}_n^{P}$	$\widehat{BP}_n^{J}$	$\widehat{BP}_n^{SM,1}$	$\widehat{BP}_n^{SM,2}$
Mean	70.76	70.77	75.75	70.75	70.76	70.26	70.27	70.26	70.24	70.27
Bias (%)	0.005	0.030	0.00	0.00	0.012	0.0017	0.0117	0.00	−0.020	−0.023
Std err	0.1419	0.1462	1.421	0.1423	0.1435	0.0643	0.0652	0.0644	0.0646	0.0648
RMSE	0.1419	0.1478	1.421	0.1423	0.1437	0.0643	0.0657	0.0644	0.0662	0.0667
Median	70.75	70.76	75.74	70.75	70.75	70.26	70.26	70.26	70.24	70.27

The table reports on the true value, the mean, the bias (in percentage), the standard error, the RMSE, and the median of $\widehat{BP}_n^{ML}$, $\widehat{BP}_n^{P}$, $\widehat{BP}_n^{J}$, $\widehat{BP}_n^{SM,1}$, and $\widehat{BP}_n^{SM,2}$ in the Vasicek model for two least persistent cases obtained from simulations and price a three-year discount bond. We simulate 5000 data sets from the Vasicek model

$$dS(t) = \kappa(\mu - S(t))dt + \sigma dB(t),$$

each with 7500 daily observations. BP is defined in Equation (21).

properly centered. The results are not surprising and hence are well explained from the binding functions in Figure 7 where the lower-right corner corresponds to the range of large values of κ. Since the slopes of the two binding functions are less than unity here as the binding function moves below the 45° line, the two simulation-based methods increase the variance in this region of the parameter space.

So far it has been assumed that the market price of risk is zero, implying that the physical measure is the same as the risk-neutral measure. To understand how the price of risk affects the performance of ML and the proposed methods, we conduct a simple Monte Carlo study by using 7500 daily observations, simulated from the Vasicek model with $\kappa = 0.018$, to estimate the price of a three-year discount bond. In the study, it is assumed that the market price parameter $\lambda = -0.1$ and is known. This setup implies that the risk-neutral process is mildly explosive, consistent with the empirical results obtained in Chen and Scott (1992, 1993) and in Geyer and Pichler (1999). Table 10 reports on the means, standard errors, RMSEs, and medians of all five estimates across 5000 replications. As in the case of zero market price of risk, the two simulation-based estimators reduce both the bias and the standard error and hence the RMSE values.

3. Estimation Bias versus Discretization Bias

As discussed earlier, to perform exact ML estimation, one needs a closed-form expression for $\ell(\theta)$ and hence $\ln f(S_t|S_{t-1}; \theta)$. Although both the Black-Scholes and the Vasicek models enable exact ML estimation, it is generally difficult to obtain a closed-form expression for $\ln f(S_t|S_{t-1}; \theta)$ in other interesting models. As a result, ML estimation may require numerical techniques or

Table 10
Finite-sample properties of $\hat{P}_n^{ML}$, $\widehat{BP}_n^P$, $\widehat{BP}_n^J$, $\widehat{BP}_n^{SM,1}$, and $\widehat{BP}_n^{SM,2}$ in the Vasicek model with nonzero market price of risk

Estimation of price of a discount bond
True value $P = 88.5705$

Estimators	$\hat{P}_n^{ML}$	$\widehat{BP}_n^P$	$\widehat{BP}_n^J$	$\hat{P}_n^{SM,1}$	$\hat{P}_n^{SM,2}$
Mean	87.46	88.13	87.86	88.32	87.94
Bias (in %)	−1.25	.495	−.807	−.283	−.716
Std error	2.093	1.987	2.273	1.930	1.874
RMSE	2.368	2.034	2.383	1.947	1.977
Median	87.99	88.94	88.29	88.92	88.57

The table reports on the true value, the mean, the bias (in percentage), the standard error, the RMSE, and the median of $\hat{P}_n^{ML}$, $\widehat{BP}_n^P$, $\widehat{BP}_n^J$, $\widehat{BP}_n^{SM,1}$, and $\widehat{BP}_n^{SM,2}$ in the Vasicek model with nonzero market price of risk ($\lambda = -0.1$), obtained from simulations. We simulate 5000 data sets from the Vasicek model

$$dS(t) = \kappa(\mu - S(t))dt + \sigma dB(t),$$

each with 7500 daily observations and price a three-year discount bond. The three-year discount bond is priced under the assumption that the market price of risk, λ, is -0.1.

simulation-based approximants. Some important work in this area of research includes Lo (1988) and Brandt and Santa-Clara (2002). Recently, methods via analytic approximations have been developed, facilitating approximation to $\ln f(S_t|S_{t-1}; \theta)$ with high accuracy and hence ML estimation. Important contributions along this line of research can be found in Aït-Sahalia (1999, 2002, 2007). We refer to Phillips and Yu (2008) for a recent discussion of alternative techniques, and briefly review here one approximate ML method that is relevant to the present study.

The Euler scheme approximates Equation (1) by the following discrete-time model:

$$S_t = S_{t-1} + \mu(S_{t-1}, \theta)h + \sigma(S_{t-1}, \theta)\sqrt{h}\epsilon_t, \tag{25}$$

where $\epsilon_t \sim$ i.i.d. $N(0, 1)$. The conditional density $f(S_t|S_{t-1})$ for the Euler discrete-time model has the following closed-form expression:

$$S_t|S_{t-1} \sim N(S_{t-1} + \mu(S_{t-1}, \theta)h, \sigma^2(S_{t-1}, \theta)h). \tag{26}$$

Denote as $\hat{\theta}_n^{AML}$ the resultant ML estimator of θ and set $\hat{P}_n^{AML} = P(\hat{\theta}_n^{AML})$. The advantage of the Euler scheme is that no matter how complicated the functions $\mu(S_t, \theta)$ and $\sigma(S_t, \theta)$ are, the conditional density and hence the log-likelihood function for the approximate model have closed-form expressions. The drawback is that the Euler scheme obviously introduces a discretization bias. The magnitude of the discretization bias depends on the size of the sampling interval h. The larger the h, the larger the discretization bias. Another disadvantage is the presence of the finite-sample estimation bias, as discussed in the last section. Both biases translate to $\hat{P}_n^{AML}$.

Simulation-based methods can be used to deal with these two types of bias simultaneously. The idea is as follows. Given a price choice p, we calculate

the implied parameter $\theta(p)$ using the pricing formula. Then we apply the Euler scheme with a much smaller step size than h (say $\delta = h/10$), which leads to the generating scheme

$$\tilde{S}_t = \tilde{S}_{t-1} + \mu(\tilde{S}_{t-1}, \theta)h + \sigma(\tilde{S}_{t-1}, \theta)\sqrt{\delta}\epsilon_t,$$

where

$$t = 0, \ldots, \frac{h}{\delta}n(= T).$$

This sequence may be regarded as a nearly exact simulation from the continuous-time model (1) with the step size δ since δ is so small. We then choose every $(1/\delta)$th observation to form the sequence of $\{\tilde{S}_t^k(p)\}_{i=1}^n$, which can be regarded as data simulated directly from model (1) with the (observationally relevant) step size h and hence these data may be regarded as having negligible discretization bias since δ is so small.

Let $\tilde{\mathbf{S}}^k(p) = \{\tilde{S}_1^k(p), \ldots, \tilde{S}_n^k(p)\}$ be data simulated from the true model, where $k = 1, \ldots, K$ with K being the number of simulated paths. Again it is important to choose the number of observations in $\tilde{\mathbf{S}}^k(p)$ to be the same as the number of observations in the observed sequence $\mathbf{S}$ for the purpose of reducing the finite-sample estimation bias. Denote by $\tilde{\phi}_n^{AML,k}(p)$ the approximate ML estimator of θ obtained from the conditional density (26) and $\tilde{\mathbf{S}}^k(p)$, and define $\tilde{P}_n^{AML,k}(p) = P(\tilde{\phi}_n^{AML,k}(p))$. The simulation-based estimation then matches $\hat{P}_n^{AML}$ with the average behavior of $\tilde{P}_n^{AML,k}(p)$. Define the average behavior by $b_n(p)$. The simulation-based estimator of p is then defined as

$$\hat{P}_n^{SM} = \operatorname{argmin}_p \left\| \hat{P}_n^{AML} - b_T(p) \right\|, \tag{27}$$

where $\| \cdot \|$ is some finite-dimensional distance metric. In practice, since $b_n(p)$ does not have an analytical expression, we calculate it as before via simulation. That is, the simulation-based estimator of p is calculated as

$$\hat{P}_{n,K}^{SM} = \operatorname{argmin}_p \left\| \hat{P}_n^{AML} - \hat{b}_{n,K}(p) \right\|, \tag{28}$$

where $\hat{b}_{n,K}(p)$ can be the sample mean or the 50th sample quantile of $\{\tilde{P}_n^{AML,k}(p)\}_{k=1}^K$.

While the above "in-fill" approach can reduce the discretization bias, there are other methods that can more effectively reduce the discretization bias. One important technique involves the use of closed-form approximations recently developed in Aït-Sahalia (1999, 2002, 2007). Jessen and Poulsen (2002) examined the relative performance of alternative techniques and found that the approach based on closed-form approximations performs the best in terms of accuracy and speed.

4. Empirical Illustrations

As an illustrative example, we now test the Vasicek model using the proposed theory based on real monthly time-series data on a short-term interest rate and real cross-section data on three discount bonds. We take the short-term interest rate as the annualized discount rates on the 13-week Treasury bill. These bills are issued by the U.S. Treasury in auctions conducted weekly by the Federal Reserve Bank. The data are reported throughout the trading day by Telerate Systems Incorporated and downloadable at finance.yahoo.com. The sample period is from March 1, 1974, to August 1, 2006, and has 390 monthly observations. The Vasicek model is used to price three discount bonds, with maturities of 5, 10, and 30 years. We obtain the yields and hence the prices of the 5-year Treasury note, the 10-year Treasury note, and the 30-year Treasury bond quoted in August 2006 in the *Wall Street Journal* (4.97%, 4.98%, and 5.07%).

To obtain the ML estimate of the theoretical price, we fit the time-series data using ML and then insert the ML estimates of the model parameters into the price formula of the zero-coupon bond. Using the asymtotic variance, we compute the 95% confidence intervals for the zero-coupon bonds. Similarly, two simulation-based methods are used to estimate the zero-coupon bonds and to obtain the corresponding 95% confidence intervals. Such a long-span series for the short-term rate is chosen because our intention is to examine the difference between the simulation approaches and the standard method in a large sample.

The time-series plot of the 13-week Treasury bill series is provided in Figure 8. Table 11 shows the sample size, mean, standard deviation, the first seven sample autocorrelations of the series, and the prices of the three discount bonds. Clearly, the Treasury bill series is highly persistent.

Table 12 reports on the estimated prices and the corresponding 95% confidence intervals of three bonds using ML and the two simulation-based methods. Although not reported in Table 12, the ML estimates of κ, μ, and σ are 0.2166, 0.0553, and 0.1934, respectively. These estimates are compatible with those reported in the literature (e.g., Aït-Sahalia 1999; Ball and Torous 1996). In particular, the estimate of κ is close to zero and suggests that the short-term interest rate is highly persistent. As pointed out in Phillips and Yu (2005), when the true value of κ is close to zero, the ML estimate of κ is biased upward, leading to downward biased estimates of theoretical contingent-claim prices. The present results corroborate that claim. The two simulation-based estimates are always higher than the MLEs. In particular, the first simulation-based estimates are 0.29%, 1.37%, and 1.48% higher than their ML counterparts, while the second simulation-based estimates are 0.25%, 0.51%, and 1.52% higher than their ML counterparts. Even with such a large sample of data, the differences in the estimates for all three bonds are large and economically significant. These findings are consistent with the magnitudes and directions of the biases and differences
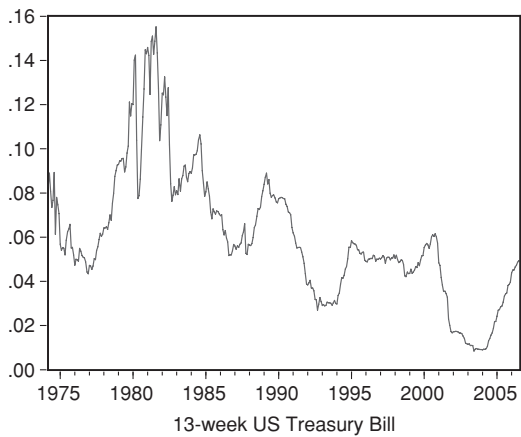
Figure 8
The dynamics of monthly 13-week Treasury bill
The graph plots the time series of monthly 13-week Treasury bill from March 1974 to August 2006.

Table 11
Summary statistics of Treasury bills and prices of Treasure notes

Time series data on Treasury bills

Number of observations	390
Mean	0.0601
Standard deviation	0.0299
Autocorrelations	
ρ_1	0.982
ρ_2	0.959
ρ_3	0.937
ρ_4	0.918
ρ_5	0.901
ρ_6	0.883
ρ_7	0.872
Price of 5-year Treasury note on Aug 1, 2006 ($\times 100$)	78.00
Price of 10-year Treasury note on Aug 1, 2006 ($\times 100$)	60.77
Price of 30-year Treasury note on Aug 1, 2006 ($\times 100$)	21.85

The 13-week Treasury bills were downloaded at finance.yahoo.com. The sample period is from March 1, 1974, to August 1, 2006, and has 390 monthly observations. We obtain the yields and hence the prices of the 5-year Treasury note, the 10-year Treasury note, and the 30-year Treasury bond quoted in August 2006 in the *Wall Street Journal* (4.97%, 4.98%, and 5.07%).

Table 12
Empirical estimates and confidence interval

	Empirical estimation of prices of discount bonds		
Maturity	5-year	10-year	30-year
Observed price	78.00	60.77	21.85
ML	77.58	59.97	21.46
95% Confidence interval	(77.46,77.70)	(59.28,60.65)	(19.80,23.13)
Simulation method 1	77.90	60.79	21.78
95% Confidence interval	(77.78,78.02)	(60.10,61.48)	(20.11,23.45)
Simulation method 2	77.87	60.27	21.79
95% Confidence interval	(77.75,78.00)	(59.59,60.96)	(20.12,23.45)

The Vasicek model is used to price three discount bonds, with maturities of 5, 10, and 30 years.

between the simulation-based and ML estimates that were found in the Monte Carlo studies.

Comparison of the observed prices with the 95% confidence intervals based on the ML estimates suggests that the data are inconsistent with a Vasicek model. In particular, out of three bonds, only one observed price (for the 30-year bond) is contained in the 95% confidence interval. On the other hand, when the observed prices are compared with the 95% confidence intervals constructed from the two simulation-based estimates, in no case can the Vasicek model be rejected. The reason for this difference is that the two simulation-based estimates are always closer to the observed counterparts. This finding suggests that bias-correction has important implications in statistical testing, particularly in the evaluation of contingent-claim pricing models.

Of course, the empirical application considered here is meant merely as an illustration of the proposed theory and not as a conclusive test of the Vasicek model. However, the empirical results do indicate that, without correcting for finite-sample bias, inferences based on ML estimation can be misleading in practically realistic cases, even when sample sizes are large.

5. Conclusions and Implications

Our findings indicate that maximum likelihood estimation, despite its many significant advantages, does not always lead to the best estimator of contingent-claim prices in finite samples in some cases. The finite-sample problem arises in the worst-case scenarios of strong nonlinearity in pricing formulas and highly persistent dynamic models.

This paper proposes two simulation-based methods to enhance the finite-sample properties of the MLE when ML is biased. The idea is based on the observation that if the MLE of a contingent-claim price is biased with actual data, then it will also be biased with simulated data. Simulations therefore enable the bias function to be calibrated for the specific model and sample size being used and from this calibrated function a bias reduction procedure is constructed that leads directly to a new simulation-based estimate. Monte Carlo studies show that, when the procedure is implemented on top of ML, it reduces not only the bias but also the variance of the MLE when the Black-Scholes model is used to price a deep-out-of-money option and when the near-unit-root Vasicek model is used to price a discount bond and a bond option. However, when ML is not biased, the simulation methods yield very similar estimates to ML.

The present paper applies this simulation-based approach to price discount bonds in the context of a Vasicek model and options in the context of a Black-Scholes model. Use of these two specific models makes it possible to employ exact ML and closed-form bond pricing and options pricing formulas. These models and simulation designs permit a comparison of the proposed methods with exact ML using replicated simulated data within feasible time frameworks. However, the technique itself is quite general and can be applied in many other

contingent-claim models. One example is the GARCH option pricing model of Heston and Nandi (2000). In a recent study, Dotsus and Markellos (2007) found that the MLE of the GARCH model of Heston and Nandi involves substantial estimation biases, even when the sample size is as large as 3000. They further show that the estimation biases may translate to option pricing. More generally, the MLE of the parameters in the multifactor affine asset pricing models of Duffie and Kan (1996) may be biased (Aït-Sahalia and Kimmel 2005) and the bias can be expected to translate to prices of contingent claims. For more general asset pricing models, the dynamics of the underlying asset price may be so complicated that an exact ML is not feasible. However, our simulation-based methods can be used in conjunction with other estimation methods, including the approximate ML method of Aït-Sahalia (2002, 2008) and the two-stage ML method of Phillips and Yu (2007). Of course, for models in which exact ML is not feasible and for contingent-claim prices that do not have closed-form expressions, the simulation-based methods will inevitably be computationally more costly.

References

Aït-Sahalia, Y. 1999. Transition Densities for Interest Rate and Other Non-linear Diffusions. *Journal of Finance* 54:1361–95.

Aït-Sahalia, Y. 2002. Maximum Likelihood Estimation of Discretely Sampled Diffusion: A Closed-Form Approximation Approach. *Econometrica* 70:223–62.

Aït-Sahalia, Y. 2008. Closed-Form Likelihood Expansions for Multivariate Diffusions. *Annals of Statistics* 36:906–37.

Aït-Sahalia, Y., and R. Kimmel. 2005. Estimating Affine Multifactor Term Structure Models Using Closed-Form Likelihood Expansion. Working Paper, Princeton University.

Aït-Sahalia, Y., and R. Kimmel. 2007. Maximum Likelihood Estimation of Stochastic Volatility Models. *Journal of Financial Economics* 83:413–52.

Aït-Sahalia, Y., and J. Yu. 2006. Saddlepoint Approximations for Continuous-Time Markov Processes. *Journal of Econometrics* 134:507–51.

Andersen, T., H. Chung, and B. Sorensen. 1999. Efficient Method of Moments Estimation of a Stochastic Volatility Model: A Monte Carlo Study. *Journal of Econometrics* 91:61–87.

Andrews, D. W. K. 1993. Exactly Median-Unbiased Estimation of First Order Autoregressive/Unit Root Models. *Econometrica* 61:139–66.

Bakshi, G., C. Cao, and Z. Chen. 1997. Empirical Performance of Alternative Option Pricing Models. *Journal of Finance* 52:2003–49.

Bakshi, G., N. Ju, and H. Ou-Yang. 2006. Estimation of Continuous-Time Models with an Application to Equity Volatility Dynamics. *Journal of Financial Economics* 82:227–49.

Ball, C. A., and W. N. Torous. 1984. The Maximum Likelihood Estimation of Security Price Volatility: Theory, Evidence, and Application to Option Pricing. *Journal of Business* 57:97–112.

Ball, C. A., and W. N. Torous. 1996. Unit Roots and the Estimation of Interest Rate Dynamics. *Journal of Empirical Finance* 3:215–38.

Black, F., and M. Scholes. 1973. The Pricing of Options and Corporate Liabilities. *Journal of Political Economy* 81:637–59.

Boyle, P. P., and A. L. Ananthanarayanan. 1977. The Impact of Variance Estimation in Option Valuation Models. *Journal of Financial Economics* 5:375–87.

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.

Butler, J. S., and B. Schachter. 1986. Unbiased Estimation of the Black/Scholes Formula. *Journal of Financial Economics* 15:341–57.

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* 47:1209–27.

Chapman, D. A., and N. Pearson. 2000. Is the Short Rate Drift Actually Non-linear? *Journal of Finance* 55:355–88.

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* 5:613–36.

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.

Cox, J., J. Ingersoll, 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.

Dotsis, G., and R. N. Markellos. 2007. The Finite Sample Properties of the GARCH Option Pricing Models. *Journal of Futures Markets* 17:599–615.

Duan, J. 1995. The GARCH Option Pricing Model. *Mathematical Finance* 5:13–32.

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* 68:1343–76.

Duffie, D., and K. J. Singleton. 1993. Simulated Moments Estimation of Markov Models of Asset Prices. *Econometrica* 61:929–52.

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.

Egorov, A. E., H. Li, and D. Ng. Forthcoming. A Tale of Two Yield Curves: Modeling the Joint Term Structure of Dollar and Euro Interest Rates. *Journal of Econometrics*.

Egorov, A. E., H. Li, and Y. Xu. 2003. Maximum Likelihood Estimation of Time-Inhomogeneous Diffusions. *Journal of Econometrics* 114:107–39.

Gallant, A., and G. Tauchen. 1996. Which Moments to Match? *Econometric Theory* 12:657–81.

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* 22:107–30.

Gouriéroux, C., A. Monfort, and E. Renault. 1993. Indirect Inference. *Journal of Applied Econometrics* 8:S85–118.

Gouriéroux, C., P. C. B. Phillips, and J. Yu. Forthcoming. Indirect Inference for Dynamic Panel Models. *Journal of Econometrics*.

Gouriéroux, C., E. Renault, and N. Touzi. 2000. Calibration by Simulation for Small Sample Bias Correction. In R. S. Mariano, T. Schuermann, and M. Weeks (eds.), *Simulation-Based Inference in Econometrics: Methods and Applications*, pp. 328–58. Cambridge, United Kingdom: Cambridge University Press.

Heston, S. L. 1993. A Closed-Form Solution for Options with Stochastic Volatility, with Applications to Bond and Currency Options. *Review of Financial Studies* 6:327–43.

Heston, S. L., and S. Nandi. 2000. A Closed-Form GARCH Option Valuation Model. *Review of Financial Studies* 13:585–625.

Huber, P. J. 1967. The Behavior of Maximum Likelihood Estimates under Nonstandard Conditions. In L. M. LeCam and J. Neyman (eds.), *Proceedings of the Fifth Berkeley Symposium on Mathematics, Statistics, and Probability*, pp. 221–33. Berkeley: University of California Press.

Hull, J. 2000. *Options, Futures, and Other Derivatives*, 2d ed. Old Tappan, NJ: Prentice-Hall.

Hull, J., and A. White. 1987. The Pricing of Options on Assets with Stochastic Volatilities. *Journal of Finance* 42:281–300.

Ingersoll, J. 1976. A Theoretical and Empirical Analysis of the Dual Purpose Fund. *Journal of Financial Economics* 3:83–124.

Jamshidian, F. 1989. An Exact Bond Option Formula. *Journal of Finance* 44:205-9.

Jensen, B., and R. Poulsen. 2002. Transition Densities of Diffusion Processes: Numerical Comparison of Approximation Techniques. *Journal of Derivatives* 9:1–15.

Kim, S., N. Shephard, and S. Chib. 1998. Stochastic Volatility: Likelihood Inference and Comparison with ARCH Models. *Review of Economic Studies* 65:361–93.

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

Lo, A. W. 1986. Statistical Tests of Contingent-Claims Asset-Pricing Models. *Journal of Financial Economics* 17:143–73.

Lo, A. W. 1988. Maximum Likelihood Estimation of Generalized Itô Processes with Discretely Sampled Data. *Econometric Theory* 4:231–47.

MacKinnon, J. G., and A. A. Smith. 1998. Approximate Bias Correction in Econometrics. *Journal of Econometrics* 85:205–30.

Merton, R. C. 1973. Theory of Rational Option Pricing. *Bell Journal of Economics and Management Science* 4:141–83.

Monfardini, C. 1998. Estimating Stochastic Volatility Models through Indirect Inference. *Econometrics Journal* 1:C113–28.

Monfort, A. 1996. A Reappraisal of Misspecified Econometric Models. *Econometric Theory* 12:597–619.

Phillips, P. C. B. 1972. The Structural Estimation of Stochastic Differential Equation Systems. *Econometrica* 40:1021–41.

Phillips, P. C. B., and J. Yu. 2005. Jackknifing Bond Option Prices. *Review of Financial Studies* 18:707–42.

Phillips, P. C. B., and J. Yu. Forthcoming. A Two-Stage Realized Volatility Approach to Estimation of Diffusion Processes with Discrete Data. *Journal of Econometrics*.

Phillips, P. C. B., and J. Yu. Forthcoming. Maximum Likelihood and Gaussian Estimation of Continuous Time Models in Finance. In T. G. Andersen, R. A. Davis, J. P. Kreiss, and T. Mikosch (eds.), *Handbook of Financial Time Series*. New York: Springer.

Pierce, J. L. 1984. Did Financial Innovation Hurt the Great Monetarist Experiment? *American Economic Review* 74:392–96.

Skaug, H., and J. Yu. 2007. Automated Likelihood Based Inference for Stochastic Volatility Models. Working Paper, Sim Kee Boon Institute for Financial Economics, Singapore Management University.

Smith, A. A. 1993. Estimating Nonlinear Time-Series Models Using Simulated Vector Autoregressions. *Journal of Applied Econometrics* 8:S63–84.

Tang, C. Y., and S. X. Chen. 2007. Parameter Estimation and Bias Correction for Diffusion Processes. Working Paper, Department of Statistics, Iowa State University.

Thompson, S. 2008. Identifying Term Structure Volatility from the LIBOR-Swap Curve. *Review of Financial Studies* 21:819–54.

Vasicek, O. 1977. An Equilibrium Characterization of the Term Structure. *Journal of Financial Economics* 5:177–86.

Yu, J. 2008. Bias in the Estimation of Mean Reversion Parameter in Continuous Time Models. Working Paper, Sim Kee Boon Institute for Financial Economics, Singapore Management University.

Zehna, P. 1966. Invariance of Maximum Likelihood Estimation. *Annals of Mathematical Statistics* 37:744.