SIMULATION-BASED ESTIMATION
OF CONTINGENT-CLAIMS PRICES

By

Peter C. B. Phillips and Jun Yu

January 2007

COWLES FOUNDATION DISCUSSION PAPER NO. 1596

Simulation-based Estimation of Contingent-claims Prices[1]

Peter C. B. Phillips

Cowles Foundation for Research in Economics
Yale University

and
University of Auckland & University of York, UK

Jun Yu
Singapore Management University

November 21, 2006

Abstract

A new methodology is proposed to estimate theortical 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. The present paper proposes instead a simulation-based method that improves the finite sample performance of the ML estimator while maintaining its good asymptotic properties. The methods are implemented and evaluated here in the Black-Scholes option pricing model and in the Vasicek bond pricing model, but have wider applicability. Monte Carlo studies show that the proposed procedures achieve bias reductions over ML estimation in pricing contingent claims. The bias reductions are sometimes accompanied by reductions in variance, leading to significant overall gains in mean squared estimation error. Empirical applications to US 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.

Keywords: Bias Reduction, Bond Pricing, Indirect Inference, Option Pricing, Simulation-based Estimation.

JEL classification: C15, G12

1. Introduction

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), and Duan (1996), 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 non-linear.

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) and Aït-Sahalia (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 formulae, 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 non-linear 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 (c.f., Lo, 1986).

In spite of its generally good asymptotic properties, ML is not necessarily the best estimation method for contingent claim prices in finite samples. There are three reasons for this. First, since the price of a contingent-claim is a non-linear transformation of the system parameters, insertion of even unbiased estimators into the pricing formulae will not assure unbiased estimation of a contingent-claim price (Ingersoll, 1976). 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 typically used for pricing claims that are contingent on short term interest rates, the MLE of the system parameters may sustain substantial finite sample bias even in very large samples as shown recently in Phillips and Yu (2005); and when biased estimated parameters are inserted into the pricing formulae, the bias can be amplified in the resulting estimates of the contingent-claim price.

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 this bias in bond pricing and bond option pricing, 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 two 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 which can achieve bias reduction as well as variance reduction, thereby offering overall gains in mean squared estimation error for contingent-claim pricing. Instead of inserting a biased

2

corrected ML estimator into the pricing formulae, 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 and Dai and Singleton, 2000) and in the context of discrete time stochastic volatility models to deal with intractable likelihoods (e.g. Monfardini, 1998, and Andersen, Chung and Lund, 1999). The methods have also been utilized to correct finite sample bias in time series models (e.g. Gourieroux, et al, 2000) and in dynamic panel models (e.g. Gourieroux, et al., 2005). 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, such as those based on the Euler approximation. When the simulation-based method is used in connection with exact 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. Third, the present methods can deal with both the estimation bias and the discretization bias that arises when non-linear stochastic differential equations are estimated. Since non-linear 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 the simulation-based methods is that they are inevitably computationally
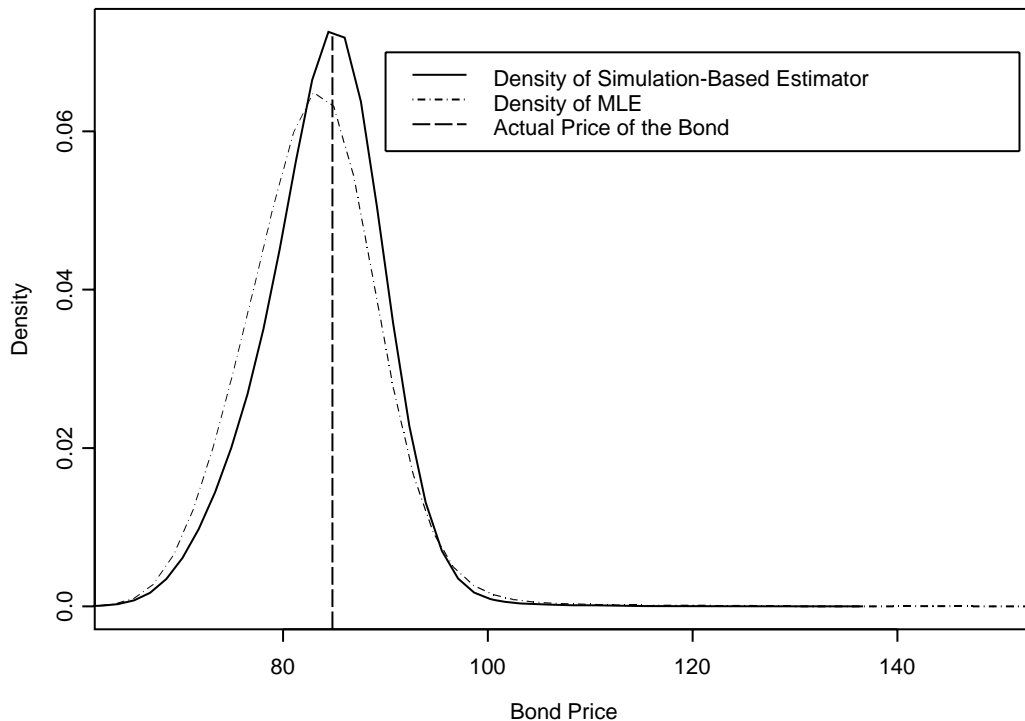
Figure 1: Distribution of simulated-based and ML estimates of bond price.

intensive. But numerical methods are now an important aspect of most empirical procedures in finance and ongoing advances in computing technology continue to make numerically intensive computations less burdensome in practical applications.

Our findings here indicate that simulation-based methods provide substantial improvements in pricing contingent-claims over ML. To illustrate, Figure 1 compares the distribution of estimates of the price of a discount bond obtained from five-year monthly data by using the MLE and a bias corrected simulation method, both in the context of a Vasicek model. The actual bond price in this case is $84.8. 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 simulated-based method is 32.4% and the gain in standard error is 24.0%. More details of this implementation and comparison are provided in Section 3.

4

The paper is organized as follows. Section 2 reviews some existing methods and introduces our simulation-based methods. Using simulated data, Section 3 illustrates the bias in 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. We also explain how the simulation-based methods can be implemented in stock option and bond valuation. The performance of these simulation-based estimates is compared with that of ML. Section 4 shows how the simulation-based methods can be used to address the estimation bias in pricing as well as the discretization bias. Section 5 examines the practical effects of simulation-based methods in an empirical application with monthly zero-coupon bond data. Section 6 concludes and outlines some further applications and implications of the approach.

2. Estimation Methods for Contingent Claims Prices

2.1. 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

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

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 a 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 any time series generating model for $S(t)$. Moreover, while we do not make a distinction between the physical measure and the risk neutral measure in the present paper, this feature can also be relaxed.

Suppose a sequence of time series observations $\mathbf{S} = (S_1, \cdots, S_T)$ taken with a sampling interval h is available and we wish to price a financial asset whose payoff is contingent upon the value of $S(t)$. 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 (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 down the log-likelihood function as

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

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 $\hat{\theta}_T^{ML}$, which is consistent, asymptotically normal and asymptotically efficient under usual regularity conditions for stationary dynamic models. In such circumstances, the limit distribution of $\hat{\theta}_T^{ML}$ is given by

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

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

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

$$\sqrt{T}(\hat{P}_T^{ML} - P) \xrightarrow{d} N(0, V_P), \tag{4}$$

where

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

Since the estimator $\hat{P}_T^{ML}$ is the MLE, it has the highest possible precision when $T \to \infty$, 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. Section 4 describes how to deal with this difficulty in the present context. Second, as shown in many studies, when the time series behavior in S_t is persistent and $\mu(S_t; \theta)$ is an affine function in S_t, (for example $\mu(S_t; \theta) = \kappa(\mu - S_t)$), the exact MLE of κ is substantially upward

---

[1] While we assume in this paper 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.

biased even in large samples (Phillips and Yu, 2005). This upward bias in $\widehat{\kappa}_T^{ML}$ translates to bias in $\hat{P}_T^{ML}$ and is large enough to be of economic significance in practice. The present paper seeks to address this problem by the use of the indirect inference procedure applied directly to P_T.

Indirect inference is a simulation-based method developed in Gourieroux, Monfort and Renault (1993) to estimate models where the likelihood is difficult to construct analytically but where the model may be readily simulated. It is closely related to the simulated GMM of Duffie and Singleton (1993). 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 $\hat{\theta}_T^{ML}$ denote the MLE that is obtained from the actual data and involves finite sample estimation bias. For any given parameter choice θ, let $\widetilde{\mathbf{S}}^k(\theta) = \{\tilde{S}_1^k, \tilde{S}_2^k, \cdots, \tilde{S}_T^k\}$ be data simulated from the time series model (1), where $k = 1, \cdots, K$ and K is the number of simulated paths. The number of observations in $\widetilde{\mathbf{S}}^k(\theta)$ is chosen to be the same as the number of actual observations in $\mathbf{S}$ so that the exact finite sample properties of $\hat{\theta}_T^{ML}$, including its finite sample bias, may be calibrated. Let $\widetilde{\phi}_T^{ML,k}(\theta)$ denote the MLE of θ obtained in this way from the k'th 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 allows for bias correction is to choose θ so that the average behavior of $\widetilde{\phi}_T^{ML,k}(\theta)$ is matched against the numerical estimate $\hat{\theta}_T^{ML}$ obtained with the observed data. In particular, the indirect inference estimator is defined by

$$\hat{\theta}_{T,K}^{II} = \operatorname{argmin}_{\theta \in \Theta} \parallel \hat{\theta}_T^{ML} - \frac{1}{K} \sum_{k=1}^{K} \widetilde{\phi}_T^{ML,k}(\theta) \parallel, \tag{6}$$

where $\parallel \cdot \parallel$ 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, $K^{-1} \sum_{k=1}^{K} \widetilde{\phi}_T^{ML,k}(\theta) \to_p E(\widetilde{\phi}_T^{ML,k}(\theta))$, applies by virtue of the nature of the simulation and then the indirect inference estimator becomes

$$\hat{\theta}_T^{II} = \operatorname{argmin}_{\theta \in \Theta} \parallel \hat{\theta}_T^{ML} - b_T(\theta) \parallel, \tag{7}$$

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

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

The procedure essentially builds in a finite-sample bias correction to $\hat{\theta}_T^{ML}$, with the bias being computed directly by simulation. Any bias that occurs in $\hat{\theta}_T^{ML}$ will also be present in

the binding function $b_T(\theta)$. Hence, with the bias correction that is built into the inversion functional $\hat{\theta}_T^{II} = b_T^{-1}(\hat{\theta}_T^{ML})$, the estimator $\hat{\theta}_T^{II}$ becomes exactly "b_T-mean-unbiased" for θ. That is, $E(b_T(\hat{\theta}_T^{II})) = b_T(\theta)$. In typical cases where $\lim_{T\to\infty} E(\hat{\theta}_T^{ML}) = \theta$ and $\hat{\theta}_T^{ML}$ is asymptotically unbiased, we have $\hat{\theta}_T^{II} \sim \hat{\theta}_T^{ML}$ in the limit as $T \to \infty$. Then, the indirect inference estimator is asymptotically equivalent to the MLE so that $\hat{\theta}_T^{II}$ shares the same good asymptotic properties of $\hat{\theta}_T^{ML}$, while having improved finite sample performance.

2.2. Direct Simulation-based Methods of Pricing

While the indirect inference estimator of θ, $\hat{\theta}_T^{II}$, may have better finite sample properties than $\hat{\theta}_T^{ML}$, inserting $\hat{\theta}_T^{II}$ into $P(\theta)$ does not necessarily lead to a better estimator than $\widehat{P}_T^{ML}$ due to the non-linearity in the pricing function. To improve the finite sample properties of $\widehat{P}_T^{ML}$, we propose to apply the simulation-based methods directly in the estimation of contingent-claims prices.

We first focus on the case where θ is a scalar. As above, we denote by $\hat{\theta}_T^{ML}$ the MLE of θ that is obtained from the actual data, and write $\widehat{P}_T^{ML} = P(\hat{\theta}_T^{ML})$. $\widehat{P}_T^{ML}$ involves finite sample estimation bias due to the non-linearity of the pricing function P in θ, or the use of the biased estimate $\hat{\theta}_T^{ML}$, 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 $\widetilde{\mathbf{S}}^k(p) = \{\tilde{S}_1^k, \tilde{S}_2^k, \cdots, \tilde{S}_T^k\}$ be data simulated from the time series model (1) given $\theta(p)$, where $k = 1, \cdots, K$ with K being the number of simulated paths. As argued above, we choose the number of observations in $\widetilde{\mathbf{S}}^k(p)$ to be the same as the number of actual observations in $\mathbf{S}$ for the express purpose of finite sample bias calibration.

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

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

Whenever bias occurs in $\widehat{P}_T^{ML}$ and from whatever source, this bias will also be present in $\widetilde{P}_T^{ML,k}(p)$ for the same reasons. Hence, the procedure builds in a finite-sample bias correction

directly to correct $\hat{P}_T^{ML}$. 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 $P(\tilde{\phi}_T^{ML,k}(p))$ as the binding function. The first one is the mean, which corresponds to the indirect inference estimation approach of 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 $\hat{P}_T^{ML}$ is highly skewed, it may be preferable to use the median in this approach. In general, however, the binding function can not 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

$$\hat{P}_{T,K}^{SIM,1} = \operatorname{argmin}_p \parallel \hat{P}_T^{ML} - \frac{1}{K} \sum_{k=1}^{K} P(\tilde{\phi}_T^{ML,k}(p)) \parallel. \tag{8}$$

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

$$\hat{P}_T^{SIM,1} = \operatorname{argmin}_p \parallel \hat{P}_T^{ML} - b_{T,1}(p) \parallel. \tag{9}$$

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

$$\hat{P}_T^{SIM,1} = b_{T,1}^{-1}(\hat{P}_T^{ML}). \tag{10}$$

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

$$\hat{P}_{T,K}^{SIM,2} = \operatorname{argmin}_p \parallel \hat{P}_T^{ML} - \hat{\rho}_{0.5} P(\tilde{\phi}_T^{ML,k}(p)) \parallel, \tag{11}$$

where $\hat{\rho}_\tau$ is the τth sample quantile obtained from $\{P(\tilde{\phi}_T^{ML,1}(p)), P(\tilde{\phi}_T^{ML,2}(p)), \cdots, P(\tilde{\phi}_T^{ML,K}(p))\}$.[2] In the case where K tends to infinity, this simulation-based estimator becomes

$$\hat{P}_T^{SIM,2} = \operatorname{argmin}_p \parallel \hat{P}_T^{ML} - b_{T,2}(p) \parallel, \tag{12}$$

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

$$\hat{P}_T^{SIM,2} = b_{T,2}^{-1}(\hat{P}_T^{ML}). \tag{13}$$

---

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

Equation (10) implies that $\hat{P}_T^{SIM,1}$ is exactly "b_T-mean-unbiased" for θ in the sense that $E(b_{T,1}(\hat{P}_T^{SIM,1})) = b_{T,1}(P)$. Similarly, from equation (13) it can be shown that $\hat{P}_T^{SIM,2}$ is exactly "b_T-median-unbiased" for θ in the sense that $\rho_{0.5}(b_{T,2}(\hat{P}_T^{SIM,2})) = b_{T,2}(P)$. If $b_{T,1}(P)$ is linear in P, exact "b_T-mean-unbiasedness" implies exact mean unbiasedness, i.e., $E(\hat{P}_T^{SIM,1}) = P$. If $b_{T,2}(P)$ is strictly monotonic in P, exact "b_T-median-unbiasedness" implies exact median unbiasedness, i.e., $\rho_{0.5}(\hat{P}_T^{SIM,1}) = P$. Thus, the sufficient condition for ensuring exact mean unbiasedness of $\hat{P}_T^{SIM,1}$ is stronger than the sufficient condition for ensuring exact median unbiasedness of $\hat{P}_T^{SIM,2}$. When $\lim_{T\to\infty} E(\hat{P}_T^{ML}) = \lim_{T\to\infty} \rho_{0.5}(\hat{P}_T^{ML}) = P$ and the slopes of the functions $b_{T,1}(P)$ and $b_{T,2}(P)$ are unity as $T \to \infty$, the two simulation-based estimators $\hat{P}_T^{SIM,1}$ and $\hat{P}_T^{SIM,2}$ are asymptotically equivalent to the MLE $\hat{P}_T^{ML}$.

Applying the delta method to equations (10) and (13), we obtain for $i = 1, 2$

$$\hat{P}_T^{SIM,i} = b_{T,i}^{-1}(\hat{P}_T^{ML}) = b_{T,i}^{-1}(b_{T,i}(P) + \hat{P}_T^{ML} - b_{T,i}(P)),$$

and

$$\mathrm{Var}(\hat{P}_T^{SIM,i}) \approx \left(\frac{\partial b_{T,i}(P)}{\partial P}\right)^{-2} \mathrm{Var}(\hat{P}_T^{ML}) \approx \left(\frac{\partial b_{T,i}(P)}{\partial P}\right)^{-2} \frac{V_P}{T}, \tag{14}$$

where V_P is given in equation (5). The asymptotic approximation (14) suggests that the simulated-based estimators should inherit some the "efficiency" properties of the ML estimator. In fact, the change in the variance depends largely on $\partial b_{T,i}(P)/\partial P$, the slope of the binding function, as seen above. For $|\partial b_{T,i}(P)/\partial P| > 1$, $\hat{P}_T^{SIM,i}$ has a smaller variance than the MLE, and for $|\partial b_{T,i}(P)/\partial P| < 1$, $\hat{P}_T^{SIM,i}$ has a larger variance than the MLE.

We now consider the case where θ is an M_θ-dimensional vector. Denote by $\hat{\theta}_T^{ML}$ 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-claims prices p in order 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

$$\theta_{\min} = \arg\min_\theta \|P(\theta) - p\|, \quad \|P(\theta) - p\| = (P(\theta) - p)'(P(\theta) - p).$$

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

10

are used to obtain the simulation-based estimator of P. Since P is now multi-dimensional, Equation (14) becomes

$$\mathrm{Var}(\hat{P}_T^{SIM,i}) \approx \left(\frac{\partial b_{T,i}(P)}{\partial P'}\right)^{-1} \mathrm{Var}(\hat{P}_T^{ML}) \left(\frac{\partial b_{T,i}(P)}{\partial P}\right)^{-1} \approx \left(\frac{\partial b_{T,i}(P)}{\partial P'}\right)^{-1} \frac{V_P}{T} \left(\frac{\partial b_{T,i}(P)}{\partial P}\right)^{-1},$$
(15)

To reduce the computational 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.

3. Illustrations and Monte Carlo Evidence

This section illustrates the bias problem in the estimation of contingent-claims prices in the context of both the Black-Scholes option pricing model and the Vasicek bond 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.

3.1. Black-Scholes Option Pricing

It is well known that in the Black-Scholes model, $\hat{P}_T^{ML}$, has little bias in most cases. However, for deep out-of-the-money options with a short time-to-maturity, $\hat{P}_T^{ML}$ can be seriously biased (see for example, Boyle and Ananthanarayanan, 1977 and Lo, 1986). It is therefore of particular interest to see how bias reduction strategies work in this case.

Let $S(t)$ be the price of an underlying stock at time t, which is assumed to follow the geometric Brownian motion process

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

and let $\{S_1, \cdots, S_T\}$ is a sample of equispaced time series observations on $S(t)$ with sampling interval h. First we define the following notation:

$$
\begin{array}{lll}
X & = & \text{Strike price,} \\
\tau & = & \text{Time to maturity,} \\
r & = & \text{Interest rate,} \\
\hat{\sigma}_T^{2,ML} & = & \text{MLE of } \sigma^2 \text{ defined by } \frac{1}{T-1}\sum_{t=1}^{T-1}\left(\ln\frac{S_{t+1}}{S_t} - \frac{1}{T-1}\sum_{t=1}^{T-1}\ln\frac{S_{t+1}}{S_t}\right)^2,
\end{array}
$$

11

$$s_T^2 = \text{Sample variance of the continuously compounded returns, defined by}$$

$$\frac{1}{T-2}\sum_{t=1}^{T-1}\left(\ln\frac{S_{t+1}}{S_t} - \frac{1}{T-1}\sum_{t=1}^{T-1}\ln\frac{S_{t+1}}{S_t}\right)^2 = \frac{T-1}{T-2}\hat{\sigma}_T^{2,ML},$$

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

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

$$\Phi = \text{Cumulative distribution function of standard normal distribution,}$$

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

In the Black-Scholes option pricing formula, the only unknown quantity is σ^2. Since $\hat{\sigma}_T^{2,ML}$ is the MLE of σ^2, $\widehat{P}_T^{ML} = P(\hat{\sigma}_T^{2,ML})$ is the MLE of P, an estimator advocated in Lo (1986). Moreover, Lo (1986) showed that

$$\sqrt{T}(\hat{P}_T^{ML} - P) \xrightarrow{d} N\left(0, \frac{\tau}{2}S^2\sigma^2\phi^2(d_1)\right). \tag{16}$$

In finite samples, however, $\hat{\sigma}_T^{2,ML}$ is biased while s_T^2 is unbiased. Inserting s_T^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 $P(s_T^2)$ and showed that $P(s_T^2)$ 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 $\widehat{P}_T^{ML}$, $P(s_T^2)$, $\hat{P}_T^{SIM,1}$ and $\hat{P}_T^{SIM,2}$. Throughout the simulations, the following parameter values are used:

$$S = \$100$$
$$\tau = 10/250$$
$$r = 5\%$$
$$T = 41$$
$$h = 1/52$$

That is, we use 40 weekly stock returns to estimate the price of a European call option which matures in two weeks and obtain the estimates $\widehat{P}_T^{ML}$, $P(s_T^2)$, $\hat{P}_T^{SIM,1}$ and $\hat{P}_T^{SIM,2}$. The experiment is replicated 5,000 times to obtain the means, standard errors, root mean square errors and medians of all four estimates. For the two simulation-based estimates, we choose the number of simulated paths to be $K = 5,000$.

12

Table 1 shows the results when $X = S\exp(r\tau)$ (i.e. the at-the-money option) and $\sigma^2 = 0.4$. In this case the actual option price is \$5.0429. Several conclusions can be drawn from the results reported in the Table. First, consistent with what has been documented, we found that $\widehat{P}_T^{ML}$ has a small percentage bias (-1.96%). Moreover, it has the smallest variance among the four estimators. Second, compared with $\widehat{P}_T^{ML}$, the use of an unbiased plug-in estimator, $P(s_T^2)$, reduces the percentage bias to -0.71%. Third, this bias is further reduced by the simulation estimators $\hat{P}_T^{SIM,1}$ (0%) and $\hat{P}_T^{SIM,2}$ (0.52%). Note that $\hat{P}_T^{SIM,1}$ is exactly mean-unbiased and $\hat{P}_T^{SIM,2}$ is exactly median-unbiased. While all three methods offer bias reduction over $\widehat{P}_T^{ML}$, they also increase the variance slightly. Finally, all four estimators perform similarly in terms of RMSE.

Table 1

Estimation of At-the-money Option Price				
True Value $P =5.0429$				
Estimators	$\widehat{P}_T^{ML}$	$P(s_T^2)$	$\hat{P}_T^{SIM,1}$	$\hat{P}_T^{SIM,2}$
Mean	4.9443	5.0072	5.0429	5.0692
Percentage bias	-1.96	-0.71	0.00	0.52
Std error	0.5573	0.5644	0.5684	0.5713
RMSE	0.5660	0.5655	0.5684	0.5719
Median	4.9186	4.9812	5.0167	5.0429

To help understand the performance of the two simulation-based methods, we plot the binding functions in Figure 2. Several features are apparent in the figure. First, both binding functions are very close to the 45 degree line, suggesting that only a small amount of bias correction is needed in the two simulation-based methods in this case. Second, both binding functions are virtually linear, implying that $\hat{P}_T^{SIM,1}$ should be exactly mean unbiased and $\hat{P}_T^{SIM,2}$ should be exactly median unbiased, a result consistent with the Monte Carlo findings in Table 1. 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 $\widehat{P}_T^{ML}$. This finding also corroborates the results found in Table 1.

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 $\sigma^2 = 0.4$. Table 2 reports the means, standard errors, root mean square errors and medians, each multiplied by 1,000, of all four estimates across 5,000 replications. The actual call option price is \$1.8038.
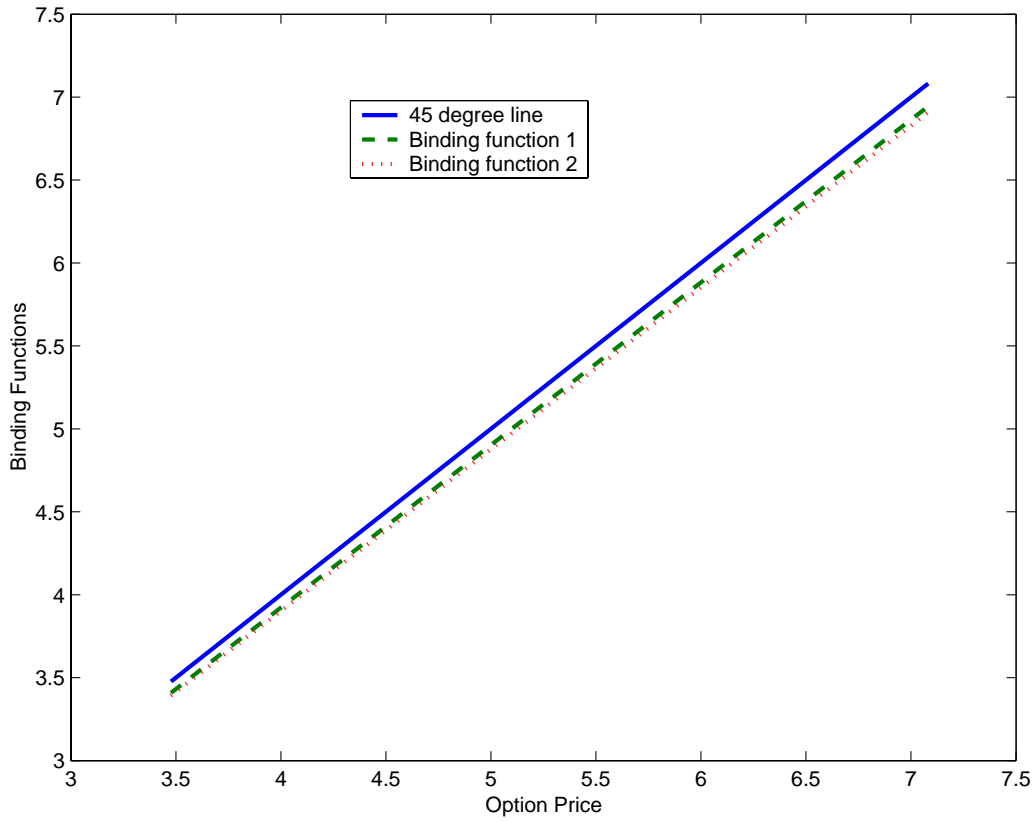
Figure 2: Binding functions of the two simulation-based methods. The 45 degree line is plotted for comparison.

Several findings emerge from Table 2. First, consistent with findings in the literature, $\widehat{P}_T^{ML}$ has a large percentage bias (21.64%). Moreover, this estimator no longer has the smallest variance. Second, compared with $\widehat{P}_T^{ML}$, instead of reducing the bias, $P(s_T^2)$ increases the percentage bias to 35.60%, so the effect of plugging in an unbiased estimator increases bias. Third and most importantly, the bias is reduced in $\hat{P}_T^{SIM,1}$ in terms of the mean and in $\hat{P}_T^{SIM,2}$ (0%) in terms of the median. The performance of $\hat{P}_T^{SIM,1}$ is particularly encouraging. This estimate not only reduces the bias in terms of the mean, but also decreases variance, producing a substantial overall gain in RMSE over $\widehat{P}_T^{ML}$ and $P(s_T^2)$. The percentage reductions in RMSE by $\widehat{P}_T^{SIM,1}$ are 21.9% and 31.1%, respectively, over $\widehat{P}_T^{ML}$ and $P(s_T^2)$. Figure 3 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 $\hat{P}_T^{SIM,2}$ 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 $\hat{P}_T^{SIM,2}$ over $\widehat{P}_T^{ML}$. The first binding function appears slightly non-linear with the slope being greater than 1 in the lower-left corner. That explains the finding that $\hat{P}_T^{SIM,1}$ does not completely remove the bias in $\widehat{P}_T^{ML}$ but can reduce the variance.

Table 2

Estimation of Deep-out-of-the-money Option Price				
True Value $P = 1.8038$				
Estimators	$\widehat{P}_T^{ML}$	$P(s_T^2)$	$\hat{P}_T^{SIM,1}$	$\hat{P}_T^{SIM,2}$
Mean	2.1941	2.4460	1.8768	2.7143
Percentage bias	21.64	35.60	4.05	50.47
Std error	2.3516	2.5677	2.222	2.7143
RMSE	2.3838	2.6468	2.2234	2.9377
Median	1.4123	1.6005	1.4036	1.8038

3.2. Vasicek Bond 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:

$$dS(t) = \kappa(\mu - S(t))dt + \sigma dB(t). \tag{17}$$

In this model, $S(t)$, the short term interest rate, mean-reverts towards the unconditional mean μ and κ measures the speed of the reversion. Define P as the price of a discount bond that pays-off \$100 at time τ. Vasicek (1977) derived the expression for the prices of a discount bond

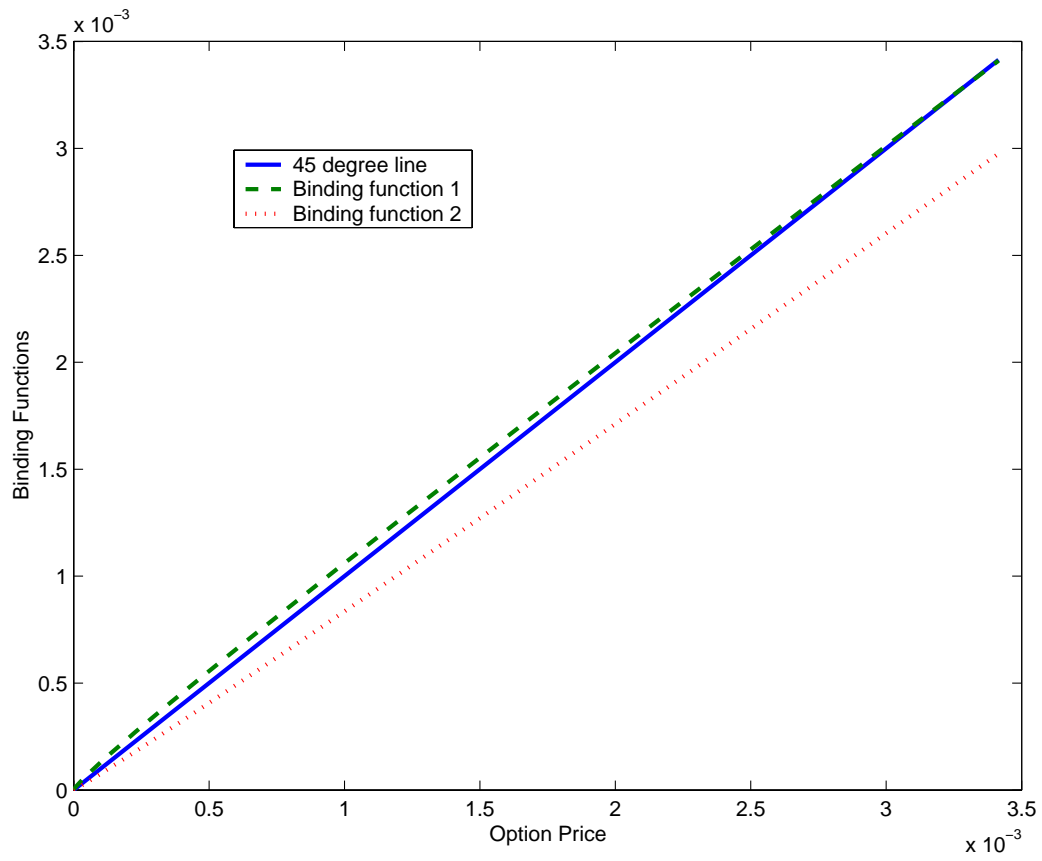$$P = 100 \times A(\tau)e^{-B(\tau)S}, \tag{18}$$

Figure 3: Binding functions of the two simulation-based methods. The 45 degree line is plotted for comparison.

where

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

$$A(\tau) = \exp\left(\frac{(B(\tau) - \tau)(\kappa^2\mu - \sigma^2/2)}{\kappa^2} - \frac{\sigma^2 B^2(\tau)}{4\kappa}\right).$$

Obviously the bond price P is a function of $\theta = (\kappa, \mu, \sigma^2)'$, the parameters in equation (17).

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

$$S_{t+1}|S_t \sim N(\mu(1 - e^{-\kappa h}) + e^{-\kappa h}S_t, \sigma^2(1 - e^{-2\kappa h})/(2\kappa)). \tag{19}$$

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 estimator of κ is severely upward biased and such a bias translates to the 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 under-estimates 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 MLE of P 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. Throughout the simulations, the following parameter values are used:

$$
\begin{aligned}
S &= 5\% \\
T &= 60 \\
\tau &= 3 \\
h &= 1/12 \\
\mu &= 0.12 \\
\sigma &= 0.01
\end{aligned}
$$

That is, we use 60 monthly observations (5 years of monthly interest rates) to estimate the price of a 3-year discount bond. Based on 60 monthly interest rates, $\widehat{P}_T^{ML}(= P(\hat{\kappa}_T^{ML}))$, $\hat{P}_T^{SIM,1}$ and $\hat{P}_T^{SIM,2}$ are all obtained. We replicate the experiment 5,000 times to obtain the means, standard errors, root mean square errors and medians of all four estimates. For the two simulation-based estimates, we chose $K = 5,000$ simulated paths.

Table 3 shows the results when the true values of κ are 0.05, 0.1, 0.15, respectively. These are empirically realistic values. The corresponding bond prices are \$84.824, \$83.677, and \$82.649.

Several conclusions can be drawn from Table 3. First, consistent with the findings in Phillips and Yu (2005), $\widehat{P}_T^{ML}$ is downward biased. In particular, the biases are -2.22%, -2.31% and -1.06% when the actual values of κ are 0.05, 0.1 and 0.15, respectively. The sizes of these estimation biases are economically significant as a 1% bias in the bond price may lead to a 25% error in an bond option price – see, for example, Hull (2000, Chapter 21.7). Moreover, $\widehat{P}_T^{ML}$ does not possess the smallest variance. On the contrary, it has the largest variance among the three estimators in all three cases. Second, the bias is reduced in $\hat{P}_T^{SIM,1}$ in terms of the mean as well as the median without any sacrifice in variance. Indeed the standard errors of $\hat{P}_T^{SIM,1}$ are 15.7%, 11.7%, 7.9% smaller than those of $\widehat{P}_T^{ML}$. In spite of these improvements, $\hat{P}_T^{SIM,1}$ still appears to underprice the bond. Thirdly, the performance of $\hat{P}_T^{SIM,2}$ is even more encouraging because both the bias and the variance (and hence the RMSE) are further reduced. Comparison of the medians with the actual bond prices suggests that $\hat{P}_T^{SIM,2}$ is median unbiased. Relative to $\widehat{P}_T^{ML}$, the biases of $\hat{P}_T^{SIM,2}$ are 24.8%, 19.8%, 15.8% smaller while standard errors of $\hat{P}_T^{SIM,2}$ are 24.0%, 21.9%, 14.2% smaller. All these gains are substantial. Figure 1 compares the nonparametric densities of $\widehat{P}_T^{ML}$ and $\hat{P}_T^{SIM,2}$ when the actual true value of $\kappa = 0.05$. 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.

Table 3

Estimators	Estimation of Price of a Discount Bond

	True Value $P = 84.824$ True Value $\kappa = 0.05$			True Value $P = 83.677$ True Value $\kappa = 0.1$			True Value $P = 82.649$ True Value $\kappa = 0.15$		
Estimators	$\widehat{P}_T^{ML}$	$\hat{P}_T^{SIM,1}$	$\hat{P}_T^{SIM,2}$	$\widehat{P}_T^{ML}$	$\hat{P}_T^{SIM,1}$	$\hat{P}_T^{SIM,2}$	$\widehat{P}_T^{ML}$	$\hat{P}_T^{SIM,1}$	$\hat{P}_T^{SIM,2}$
Mean	82.938	83.598	83.556	81.747	82.457	82.536	80.811	81.579	81.744
% bias	-1.06	-0.27	-0.32	-2.31	-1.46	-1.36	-2.22	-1.29	-1.09
Std err	5.3526	4.5149	4.0658	5.4807	4.8418	4.3947	5.3650	4.9414	4.5159
RMSE	5.6636	4.6786	4.2592	5.8103	4.9929	4.5403	5.6712	5.0560	4.6058
Median	83.268	85.056	84.824	81.676	83.601	83.677	80.594	82.397	82.649

Figure 4 plots the binding functions. Both binding functions involve some interesting non-linearity. In particular, while the binding functions are more linear in the lower-left corner, they become more non-linear in the upper-right corner. Note that the upper-right corner corresponds to the range of low values of κ, a region which is empirically more realistic. Since the second binding function is still monotonic, $\hat{P}_T^{SIM,2}$ remains median unbiased. However, the non-linearity in the first binding function implies that $\hat{P}_T^{SIM,1}$ is not mean unbiased. Moreover, in the upper-right corner, the slopes of the two binding functions are larger than 1, explaining
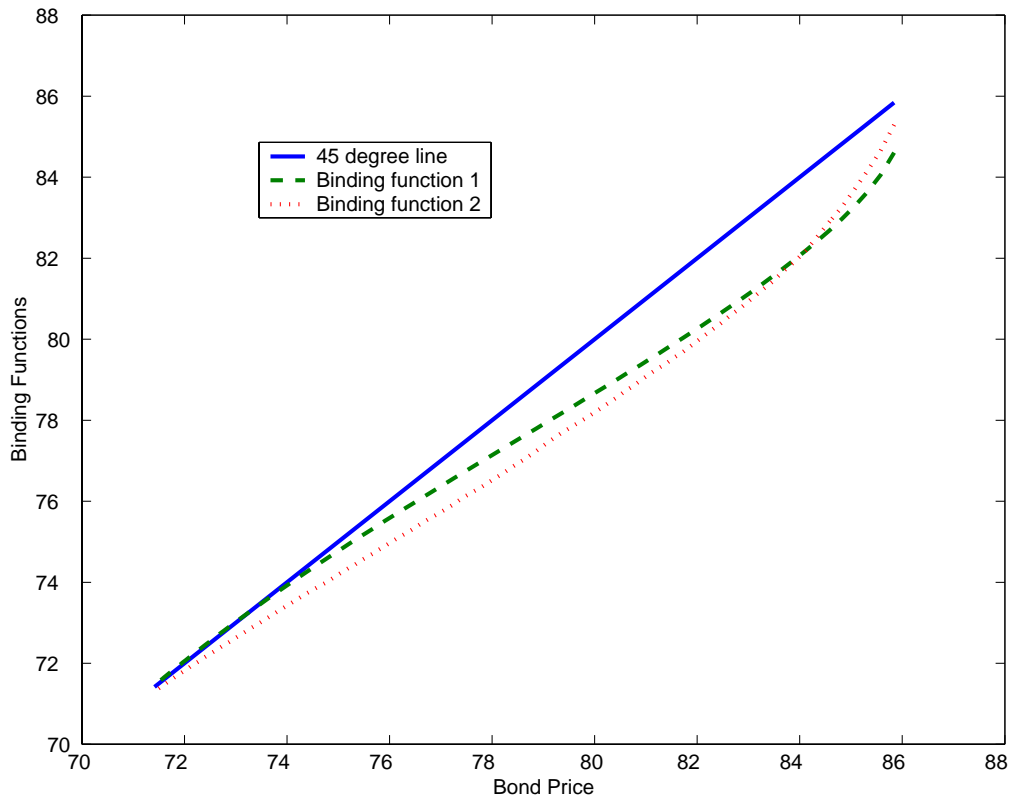
Figure 4: Binding functions of the two simulation-based methods. The 45 degree line is plotted for comparison.

why the variance is reduced in $\hat{P}_T^{SIM,1}$ and $\hat{P}_T^{SIM,2}$. All the results are confirmed in the Monte Carlo study.

4. 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 model and the Vasicek model enable exact ML estimation, it is generally impossible to obtain a closed form expression for $\ln f(S_t|S_{t-1};\theta)$ for other interesting models. As a result, ML estimation requires numerical techniques or analytic or simulation-based approximants. Some important work in this area of research include Lo (1988), Brandt and Santa-Clara (2002) and Aït-Sahalia (2002). We refer to Phillips and Yu (2006) for a recent discussion of alternative estimation 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

$$S_t = S_{t-1} + \mu(S_{t-1},\theta)h + \sigma(S_{t-1},\theta)\sqrt{h}\epsilon_t, \tag{20}$$

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:

$$S_t|S_{t-1} \sim N\left(S_{t-1} + \mu(S_{t-1},\theta)h, \sigma^2(S_{t-1},\theta)h\right). \tag{21}$$

Denote as $\widehat{\theta}_T^{AML}$ the resultant ML estimator of θ and set $\widehat{P}_T^{AML} = P(\widehat{\theta}_T^{AML})$. 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 is 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 $\widehat{P}_T^{AML}$.

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

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

where

$$t = 1,\cdots,\frac{h}{\delta}T + 1(= N).$$

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 $\{\widetilde{S}_t^k(p)\}_{i=1}^T$, which can be regarded as data simulated directly from model (1) with the (observationally relevant) step size h and hence this data may be regarded as having negligible discretization bias since δ is so small.

Let $\tilde{\mathbf{S}}^k(p) = \{\widetilde{S}_1^k(p), \cdots, \widetilde{S}_T^k(p)\}$ be data simulated from the true model, where $k = 1, \cdots, K$ 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 $\widetilde{\phi}_T^{AML,k}(p)$ the approximate ML estimator of θ obtained from the conditional density (21) and $\tilde{\mathbf{S}}^k(p)$, and define $\widetilde{P}_T^{AML,k}(p) = P(\widetilde{\phi}_T^{AML,k}(p))$. The simulation-based estimation then matches $\widehat{P}_T^{AML}$ with the average behavior of $\widetilde{P}_T^{AML,k}(p)$. Define the average behavior by $b_T(p)$. The simulation-based estimator of p is then defined as,

$$\widehat{P}_T^{SIM} = \mathrm{argmin}_p \parallel \widehat{P}_T^{AML} - b_T(p) \parallel . \tag{22}$$

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

$$\widehat{P}_{T,K}^{SIM} = \mathrm{argmin}_p \parallel \widehat{P}_T^{AML} - \widehat{b}_{T,K}(p) \parallel, \tag{23}$$

where $\widehat{b}_{T,K}(p)$ can be the sample mean or the 50th sample quantile of $\{\widetilde{P}_T^{AML,k}(p)\}_{k=1}^K$.

5. 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 Yahoo.Finance. 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 on August 1, 2006 in *the Wall Street Journal* (4.97%, 4.98% and 5.07%).

Figure 5:
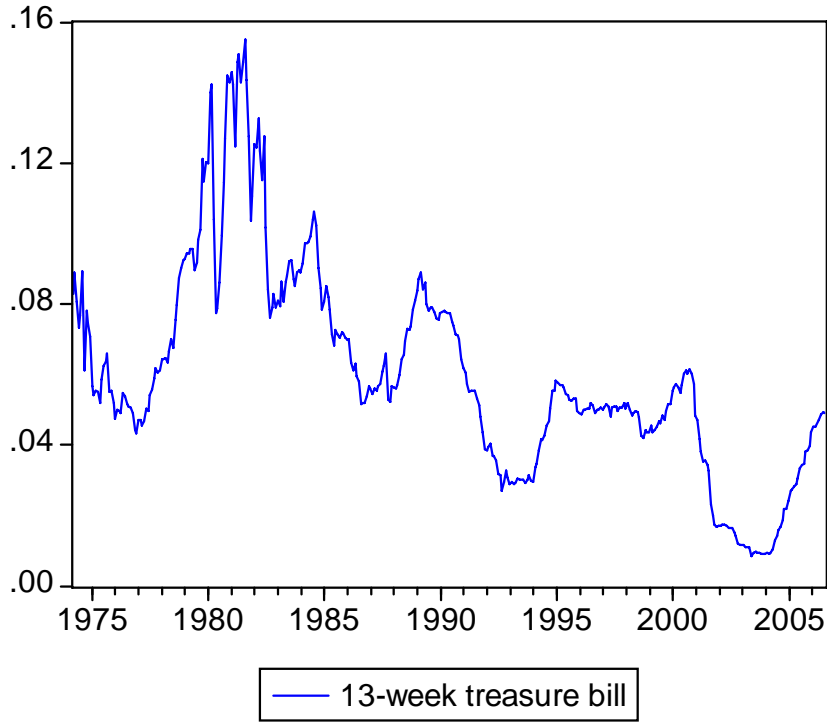
To obtain the ML estimate of the theoretical price, we fit the time series data using ML and then insert the ML estimates of the model parameters into the price formula of the zero-coupon bond. Using the asymtotic variance, we compute the 95% confidence intervals for the zero-coupon bonds. Similarly, 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 treasure bills is provided in Fig. 5. Table 4 shows the sample size, mean, standard deviation, the first seven sample autocorrelations of the series and prices of the three discount bonds. Clearly, the 13-week Treasury bills are highly persistent.

Tables 5 reports 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 5, 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) and 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 is severely biased upward, leading to 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 between the simulation-based and ML estimates that were found in the Monte Carlo studies.

Table 4: Summary Statistics

Time Series Data on Treasure Bill	
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 Treasure Note on Aug 1, 2006 ($\times 100$)	78.00
Price of 10-year Treasure Note on Aug 1, 2006 ($\times 100$)	60.77
Price of 30-year Treasure Note on Aug 1, 2006 ($\times 100$)	21.85

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 of a 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

23

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 presented above do indicate that, without correcting the finite-sample bias, inferences based on the ML estimates can be misleading in practically realistic cases, even when sample sizes are large.

<div align="center">

Table 5

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)

</div>

6. Conclusions and Implications

Our findings indicate that maximum likelihood estimation does not always lead to the best estimator of contingent-claim prices in finite samples. One reason for the shortcoming is the non-linear nature of the pricing formulae. Another explanation is the finite sample estimation bias of the MLE in dynamic models.

This paper proposes two simulation-based methods to improve the finite sample properties of the MLE. 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 the procedure 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 Vasicek model is used to price a discount bond.

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 formulae. 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 obvious example is the pricing of bond options (Jamshidian, 1990). Another is the GARCH option pricing model of Heston and Nandi (2000). In a recent study, Dotsus and Markellos (2006) 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 3,000. They further show that the estimation biases may translate to option pricing. More generally, the MLE of the parameters in the multi-factor 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 exact ML is not feasible. However, our simulation-based methods can be used in connection with other estimation methods, including the approximate ML method of Aït-Sahalia (2002) and Duffie, Pedersen and Singleton (2003). Of course, for models in which exact ML is not feasible and for contingent-claims prices which do not have closed-form expressions, the simulation-based methods will inevitably be computationally more costly.

7. References

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

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

Aït-Sahalia, Y., and R. Kimmel, 2005, "Estimating Affine Multifactor Term Structure Models Using Closed-form Likelihood Expansion," working paper, Princeton University.

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, 139–166.

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, 215–238.

Black, F. and M. Scholes, 1973, "The Pricing of Options and Corporate Liabilities," *Journal of Political Economy* 81, 637-659.

Boyle, P. P. and AL Ananthanarayanan, 1977, "The Impact of Variance Estimation in Option Valuation Models," *Journal of Financial Economics* 5, 375-387.

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, 341-357

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

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

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

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

Dotsis, G. and R.N. Markellos, 2006, "The Finite Sample Properties of the GARCH Option Pricing Models," *Journal of Futures Markets*, forthcoming.

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

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

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

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.

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

Gouriéroux, C., Phillips, P.C.B. and J. Yu, 2005, "Indirect inference for dynamic panel models," Working paper, Cowles Foundation for Research in Economics, Yale University.

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*, Cambridge University Press, p328-358.

Heston, S.L. 1993. "A closed-form solution for options with stochastic volatility, with application to bond and currency options," *Review of Financial Studies* 6, 327–343.

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," *Proceedings of the Fifth Berkeley Symposium on Mathematics, Statistics, and Probability* 1, 221-233.

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

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, 205-209.

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, 143–173.

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

Merton, R. C., 1973, "Theory of Rational Option Pricing," *Bell Journal of Economics and Management Science*, 4, 141–183.

Monfardini, C., 1998, "Estimating stochastic volatility models through indirect inference", *The Econometrics Journal* 1, C113-C128.

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, 1021-1041.

Phillips, P.C.B. and J. Yu, 2005, "Jackknifing bond option prices," *Review of Financial Studies*, 18, 707-742.

Phillips, P.C.B. and J. Yu, 2006, "Maximum Likelihood and Gaussian Estimation of Continuous Time Models in Finance," Cowles Foundation for Research in Economics, Working Paper.

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

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

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