Clifford A. Ball
Walter N. Torous
University of Michigan

The Maximum Likelihood Estimation of Security Price Volatility: Theory, Evidence, and Application to Option Pricing*

Assuming security price dynamics are governed by a diffusion process and given the publicly most available form of data, this paper provides the maximum likelihood estimator of security price volatility. A Monte Carlo simulation compares the small-sample properties of this and other proposed security price volatility estimators. The resultant maximum likelihood estimator of the Black-Scholes call option price formulation is also derived. For small sample sizes a Monte Carlo simulation study facilitates comparison with other proposed estimation procedures. Also, the statistically most powerful confidence intervals for the Black-Scholes call option price are constructed.

I. Introduction

Several statistical estimators of security price volatility have recently been put forth in the finance literature. Given public data and a posited model of security price dynamics, these alternative procedures lead to estimates of varying statistical efficiency. Assuming that security prices are governed by a diffusion process, this paper derives and subsequently examines the properties of the most efficient estimator of security price volatility, using the most available form of public data: historical high, low, and closing prices.

Efficient security price volatility estimation is of considerable importance in financial research. The seminal analysis of Black and Scholes (1973) establishes an equilibrium option pricing formula whose only unobservable input is the variance of the underlying security. Clearly, the precision of the variance estimate translates to the estimate of the corresponding option price. Also, Merton's (1980) exploratory investigation concluded that efficient estimation of the expected return on the market requires efficient estimation of the

* We would like to thank the participants of the Statistics Department Seminar and Finance Department Workshop at the University of Michigan for their helpful comments. We are especially grateful to Fred Hoppe, Adrian Tschoegl, and an anonymous referee for their insightful suggestions. Any remaining errors are our responsibility.

(*Journal of Business*, 1984, vol. 57, no. 1, pt. 1)

0021-9398/84/5701-0003$01.50

variance of the market. That the expected return on the market is an integral component of financial valuation models in itself warrants the most efficient estimation procedure. Moreover, Rosenberg (1972) and others have empirically established the nonstationarity of security price volatilities. As such, the most efficient estimation procedure is to be preferred, since it is least susceptible to these nonstationarities. That is, any predesignated degree of statistical accuracy can be achieved with the least amount of historical data.

A number of statistical procedures exist to estimate volatility, given that security prices are governed by a diffusion process and estimators are based on publicly available data. The classical approach considers only closing prices, and an unbiased estimator of variance is derived by taking an average of the square of corresponding closing price differences. Unfortunately, this approach excludes other publicly available information. Parkinson (1980) establishes an alternative and statistically more efficient security price volatility estimator based on the difference between daily high and low prices. Garman and Klass (1980) improve on this procedure by exploiting the interdependence of high, low, and closing prices. Requiring their estimator to be analytic in a neighborhood of the origin and unbiased, their analysis relies on the first few moments of the joint probability distribution of daily high, low, and closing prices. However, considering the entire probability distribution as opposed only to the first few moments must provide additional information and lead to a statistically preferred estimation procedure. Consequently, this paper derives the maximum likelihood estimator of security price volatility based on daily high, low, and closing prices.

Since a substantial application of security price volatility estimation is option price estimation, the maximum likelihood estimator of the Black-Scholes call option price is subsequently derived. The resultant estimator possesses the highest possible degree of statistical efficiency. Furthermore, its asymptotic properties allow the formulation of the shortest (on average) and thus the statistically most powerful confidence intervals for the Black-Scholes call option price.

In estimating security price volatility, practitioners may not have available a sufficiently large time series of closing, high, and low security prices. More important, given the nonstationarity of security price volatilities, practitioners may wish to restrict their attention to a small sample of data. In the light of this empirical limitation, the small-sample properties of the proposed security price volatility estimators are also explored.

The plan of this paper is as follows: In Section II the maximum likelihood estimator of security price volatility and its asymptotic properties are derived. The small-sample properties of the classical, Parkinson, Garman and Klass, and maximum likelihood estimators are investigated in Section III. The error properties of these volatility estimators are not well known and a controlled simulation study will remedy this deficiency. These results can facilitate practitioners' choice amongst these alternative estimators. In Section IV the error properties of the corresponding Black-Scholes call option price estimators are determined, and optimal confidence intervals are constructed. The comparable confidence intervals of Boyle and Ananthanarayanan (1977) are presented to highlight our results. Section V provides our conclusions.

II. Maximum Likelihood Estimation

The dynamics of security prices will be assumed governed by

$$\ln P(t) = \sigma Z(t) \equiv X(t), \tag{1}$$

where $P(t)$ is the security price at time t, σ the diffusion coefficient to be estimated, and $Z(t)$ is a standardized Wiener process. Therefore, the logarithm of security prices is a nonstandardized Wiener process.

Given the posited model of security price dynamics and the historical data most accessible to investors—a time series of daily closing, high, and low prices—we shall derive an optimal procedure for estimating the diffusion coefficient σ. The maximum likelihood procedure is such an optimal procedure. It has intuitive appeal since the resultant maximum likelihood estimate is that estimate for which the underlying data were most likely. Moreover, the procedure is asymptotically optimal under rather mild regularity conditions (see Cramer 1946). In particular, the maximum likelihood estimator is asymptotically unbiased, consistent, and efficient, in that the variance of the estimator attains its Cramer-Rao lower bound. These properties make the maximum likelihood estimation of the diffusion coefficient σ rather desirable.

The likelihood function of σ, or equivalently σ^2, estimated from a time series of daily closing, high, and low security prices, equals the probability density of the given time series conditional on σ^2. Defining

$$A(t) = \max_{0 \leq \tau \leq t} X(\tau) \equiv \text{high}$$

and

$$-B(t) = \min_{0 \leq \tau \leq t} X(\tau) \equiv \text{low},$$

it follows that this joint density is given by

$$f(x, a, b; \sigma^2) = \frac{\partial^3}{\partial x \partial a \partial b} P[X(t) \leq x, A(t) \leq a, B(t) \leq b]$$

where $a > 0$, $b > 0$, and $-b \leq x \leq a$. Here and throughout the closing, high, and low security prices are assumed normalized relative to the opening security price. However, note that by definition

$$\frac{\partial}{\partial x} P[X(t) \leq x, A(t) \leq a, B(t) \leq b]$$
$$= \frac{\partial}{\partial x} P[X(t) \leq x, -b \leq X(\tau) \leq a \quad \text{for } 0 \leq \tau \leq t]$$

is the transition density for a nonstandardized Wiener process with absorbing barriers at a and $-b$. This transition density, $p(x, t)$, is the solution to the Kolmogorov forward equation

$$\frac{1}{2}\sigma^2 \frac{\partial^2 p}{\partial x^2} = \frac{\partial p}{\partial t},$$

subject to the boundary conditions $p(a, t) = 0$ and $p(-b, t) = 0$ (see Cox and Miller 1972, esp. p. 220). This is the heat equation with solution

$$p(x, t) = \sum_{n=1}^{\infty} \left[\frac{2}{a + b} \sin\left(\frac{n\pi b}{a + b}\right)\right] \sin\left[\frac{n\pi(x + b)}{a + b}\right] \exp\left[-\frac{n^2\pi^2\sigma^2 t}{2(a + b)^2}\right].$$

This series is absolutely convergent and differentiable term by term. Without loss of generality, we set $t = 1$, then differentiate with respect to a and b, and the joint density of the daily closing, high, and low security prices conditional on σ^2 follows:

$$f(x, a, b; \theta) = \sum_{n=1}^{\infty} \exp[-\delta(n)\theta][\alpha(n)\theta^2 + \beta(n)\theta + \gamma(n)], \quad (2)$$

where $\theta = \sigma^2$ and $\alpha(n)$, $\beta(n)$, $\gamma(n)$, and $\delta(n)$ are as displayed in Appendix A.

Assuming a time series of M independent observations on the daily closing, high, and low security price, the logarithm of the likelihood function of θ is

$$\ln L(\theta) = \sum_{i=1}^{M} \ln f(x_i, a_i, b_i; \theta).$$

Derivatives of the likelihood function for all relevant orders exist. Therefore the regularity conditions of Cramer (1946) are satisfied. A necessary condition for the existence of the maximum likelihood estimate is $[d\ln L(\theta)]/(d\theta) = 0$. Given the joint density function (2), the first-order condition is nonlinear in the parameter, implying that recourse must be had to numerical procedures. Appendix B outlines the appropriate solution procedure. Suffice it to say that given an appropriate starting value (e.g., the Garman and Klass estimate), we consistently achieved convergence to a unique maximum in no more than four iterations. As a consequence, the computational effort required to implement the maximum likelihood algorithm is minimal.

The maximum likelihood estimator is asymptotically normally distributed, with mean θ_0, the true value of the security price volatility, and variance given by

$$\frac{-1}{E_{\theta_0}\left(\frac{d^2 \ln L(\theta)}{d\theta^2}\right)},$$

the Cramer-Rao lower bound. In other words, if it is assumed that security prices are governed by a diffusion process, (1), given the publicly most available data the maximum likelihood estimator is the minimum variance bound (MVB) estimator of security price volatility. The asymptotic normality of the maximum likelihood estimator also allows the construction of confidence intervals. In particular, an approximate $(1 - \alpha) \times 100\%$ confidence interval for θ is

$$\hat{\theta} \pm Z_{\alpha/2} \sqrt{-\left(\frac{d^2 \ln L(\theta)}{d\theta^2}\right)^{-1}_{\theta = \hat{\theta}}},$$

where $\hat{\theta}$ is the maximum likelihood estimate and $Z_{\alpha/2}$ is the $(1 - \alpha/2) \times 100$th percentile of the standard normal distribution.

The maximum likelihood procedure has a number of advantageous properties, at least for large samples. Yet practitioners may not have access to a sufficiently large sample of daily closing, high, and low security prices. Indeed, given the nonstationarity of security price volatilities, practitioners may not wish to avail themselves of a sufficiently large sample. An investigation, then, regarding the small-sample properties of the maximum likelihood estimator is warranted. The small sample properties of other security price volatility estimators will also be given. Such information will facilitate a choice amongst the alternative estimators.

III. Small-Sample Properties of Volatility Estimators

The small-sample properties of the maximum likelihood security price volatility estimator and other proposed security price volatility estimators were derived by a Monte Carlo simulation. In particular, security prices were simulated on the assumption that their dynamics were governed by the diffusion process (1), with $\theta \equiv \sigma^2 = .1$. The alternative estimators were computed assuming samples of $M = 5, 10, 20, 40$, and 60 days' data. To derive corresponding sample distributions, these experiments were repeated 1,000 times. For further details regarding the mechanics of the simulation, the reader is referred to Appendix C.

Employing the current notation and assuming, without loss of generality, a single day's data, we may summarize the alternative security price volatility estimators as follows. The classical estimator based solely on closing prices is given by $\hat{\theta}_1 = x^2$. Parkinson's estimator is

$$\hat{\theta}_2 = \frac{(a - b)^2}{4 \ln 2},$$

while the Garman and Klass estimator is

$$\hat{\theta}_3 = .511(a - b)^2 - .019[x(a + b) - 2ab] - .383x^2.$$

The maximum likelihood estimator will be denoted by $\hat{\theta}_4$ throughout. The sample distributions of these alternative estimators will be summarized by statistics based on their first four moments: mean, variance, skewness, and kurtosis. Their relative statistical efficiency will also be determined. The relative statistical efficiency of an estimator, $\hat{\theta}_i$, is defined by the ratio

$$\frac{\text{var}(\hat{\theta}_1)}{\text{var}(\hat{\theta}_i)}.$$

The larger this ratio the more efficient the estimation procedure, in that a correspondingly smaller number of data are required to achieve any given degree of statistical accuracy. Table 1 summarizes our results. For moderate sample sizes ($M \geq 20$) the maximum likelihood estimator

TABLE 1 Security Volatility Estimation Results

Sample Size	Statistic	Mean	Variance	Skew	Kurtosis	Sample Relative Efficiency
5	$\hat{\theta}_1$	.098	38.88×10^{-4}	1.20	2.20	1
	$\hat{\theta}_2$	.099	7.35×10^{-4}	.78	.82	5.29
	$\hat{\theta}_3$	.099	4.92×10^{-4}	.61	.48	7.90
	$\hat{\theta}_4$	.101	4.72×10^{-4}	.63	.66	8.24
10	$\hat{\theta}_1$	.101	20.59×10^{-4}	1.12	3.39	1
	$\hat{\theta}_2$	.099	4.25×10^{-4}	.68	.93	4.85
	$\hat{\theta}_3$	.099	2.71×10^{-4}	.38	.09	7.60
	$\hat{\theta}_4$	.099	2.49×10^{-4}	.50	.30	8.27
20	$\hat{\theta}_1$	.100	9.16×10^{-4}	.54	.49	1
	$\hat{\theta}_2$	.099	1.97×10^{-4}	.41	.90	4.65
	$\hat{\theta}_3$	.099	1.38×10^{-4}	.34	.40	6.64
	$\hat{\theta}_4$	.100	1.19×10^{-4}	.30	.51	7.70
40	$\hat{\theta}_1$	.100	4.78×10^{-4}	.43	.03	1
	$\hat{\theta}_2$	.100	$.98 \times 10^{-4}$	.38	.22	4.88
	$\hat{\theta}_3$	.099	$.71 \times 10^{-4}$	.31	.36	6.73
	$\hat{\theta}_4$	.100	$.62 \times 10^{-4}$	.32	.06	7.71
60	$\hat{\theta}_1$	.099	3.05×10^{-4}	.27	.11	1
	$\hat{\theta}_2$	.100	$.61 \times 10^{-4}$	.22	.29	5.00
	$\hat{\theta}_3$	.100	$.43 \times 10^{-4}$	.14	.31	7.09
	$\hat{\theta}_4$	.100	$.35 \times 10^{-4}$	.06	.01	8.71

is effectively unbiased. Furthermore, for all sample sizes it possesses minimum variance.

Implicit in the preceding analysis is the assumption that security price dynamics over the period of time when markets are open are identical to security price dynamics over the period of time when markets are closed. Consequently, incorporating opening prices will improve the relative statistical efficiency of all the proposed estimators. Therefore, following Garman and Klass (1980), composite estimators may be formed:

$$\hat{\theta}_i^* = \omega_i \frac{O_p^2}{f} + (1 - \omega_i) \frac{\hat{\theta}_i}{(1 - f)} \quad i = 1, 2, 3, 4,$$

where again, without loss of generality, a single day's worth of data is assumed. Here, f denotes the fraction of the day that trading is closed, O_p denotes the normalized opening price, while the weight ω_i is chosen so that $\text{var}(\hat{\theta}_i^*)$ is minimized. By inspection, the choice of ω_i is independent of f and, additionally, we have the following:

LEMMA: The variance of the composite estimator $\hat{\theta}_i^*$ is minimized for $\omega_i = (1 + E_i)^{-1}$, where E_i is the relative statistical efficiency of the estimator $\hat{\theta}_i$. Furthermore, the relative statistical efficiency of the composite estimator $\hat{\theta}_i^*$ is given by $E_i^* = E_i + 1$.

PROOF: Without loss of generality assume

$$\hat{\theta}_i^* = \omega_i O_p^2 + (1 - \omega_i) \hat{\theta}_i \quad i = 1, 2, 3, 4.$$

A Wiener process possesses independent increments, implying that

$$\begin{aligned} \text{var}(\hat{\theta}_i^*) &= \omega_i^2 \text{var}(O_p^2) + (1 - \omega_i)^2 \text{var}(\hat{\theta}_i) \\ &= \omega_i^2 (1 + E_i) E_i^{-1} - 2\omega_i E_i^{-1} + E_i^{-1} \quad i = 1, 2, 3, 4. \end{aligned}$$

It immediately follows that $\text{var}(\hat{\theta}_i^*)$ is minimized for $\omega_i = (1 + E_i)^{-1}$ and the corresponding minimum value is given by $(1 + E_i)^{-1}$. Q.E.D.

An implication of the lemma is that the relative statistical efficiency of the composite maximum likelihood estimator remains highest.

The simulation results indicate the desirability of the maximum likelihood estimator. Whereas Garman and Klass restrict their estimator to be both analytic in a neighborhood of the origin and unbiased, if we relax these requirements the maximum likelihood estimator yields increased efficiency while maintaining practical unbiasedness.

IV. Call Option Price Estimation

A significant application of security price volatility estimation is option price estimation. Black and Scholes (1973) derive an equilibrium option price whose only unobservable (and hence to be estimated) input is the variance per period of the rate of return on the underlying stock. Letting ψ denote the price of the call option, we have $\psi \equiv \psi(\theta) = PN(d_1) - E\exp(-rT)N(d_2)$, where

$P \equiv$ current price of the stock,
$E \equiv$ exercise price of the option,
$r \equiv$ continuously compounded riskless rate,
$T \equiv$ time to maturity of the option,
$N(\cdot) \equiv$ cumulative standard normal distribution function,
$d_1 = [\ln(P/E) + (r + \theta/2)T]/\sqrt{\theta T}$,
$d_2 = d_1 - \sqrt{\theta T}$, and
$\theta \equiv$ as before, variance per period of the rate of return on the underlying stock.

Given an estimate of θ, an estimate of the call option price $\psi(\theta)$ follows. Notice that the call option price is a nonlinear but one-to-one function of the variance per period of the rate of return on the underlying stock.

An attractive property of the maximum likelihood estimation procedure is its invariance under one-to-one parameter transformations: If $\hat{\theta}$ is the maximum likelihood estimator of θ and $\psi(\theta)$ is any one-to-one function of θ, $\psi(\hat{\theta})$ is the maximum likelihood estimator of $\psi(\theta)$. It follows, given the maximum likelihood estimate of security price volatility based on a time series of daily closing, high, and low security prices (i.e., $\hat{\theta}_4$), that the maximum likelihood estimator of the call option price based on this data is simply $\psi(\hat{\theta}_4)$, where ψ is the Black-Scholes call option formulation. All the desirable asymptotic properties of the maximum likelihood estimation procedure are preserved under this one-to-one transformation. However, it should be emphasized that the Black-Scholes formulation is a nonlinear transformation of the variance of the underlying stock. Therefore, the small sample properties of the proposed variance estimators may not be preserved under this nonlinear transformation. To investigate the corresponding small sample properties, a Monte Carlo simulation was replicated for the alternative call option price estimators.

For this simulation out-, at-, and in-the-money call options are considered. Specifically, we assume $P = \$40$, $r = .10$, $T = .25$, $\theta = .1$ and $E = \$35$, $\$40$, and $\$45$, yielding call option prices of \$.85, \$3.01, and \$6.64, respectively. The results are presented in tables 2–4. For all sample sizes and across all option contracts the maximum likelihood estimator possesses maximum relative efficiency. However, the relative efficiency of all estimation procedures is maximized for at-the-money options. Unbiasedness of variance estimators is not preserved under the nonlinear Black-Scholes call option price transformation (Boyle and Ananthanarayanan 1977). For small sample sizes we detect this bias, but for moderate sample sizes ($M \geq 20$) this effect is not discernable.

TABLE 2 Out-of-the-Money Call Option Simulation Results

Sample Size	Statistic	Mean	Variance	Skew	Kurtosis	Sample Relative Efficiency
5	$\psi(\hat{\theta}_1)$	.79	28.49×10^{-2}	.78	.48	1
	$\psi(\hat{\theta}_2)$	.83	5.97×10^{-2}	.57	.32	4.77
	$\psi(\hat{\theta}_3)$	.84	4.05×10^{-2}	.43	.18	7.03
	$\psi(\hat{\theta}_4)$	.86	3.85×10^{-2}	.45	.33	7.40
10	$\psi(\hat{\theta}_1)$	.84	15.71×10^{-2}	.65	.91	1
	$\psi(\hat{\theta}_2)$	.84	3.48×10^{-2}	.50	.45	4.51
	$\psi(\hat{\theta}_3)$	.84	2.27×10^{-2}	.26	−.18	6.92
	$\psi(\hat{\theta}_4)$	.84	2.08×10^{-2}	.36	.30	7.55
20	$\psi(\hat{\theta}_1)$	.84	7.47×10^{-2}	.30	.05	1
	$\psi(\hat{\theta}_2)$	.85	1.65×10^{-2}	.27	.54	4.53
	$\psi(\hat{\theta}_3)$	.84	1.16×10^{-2}	.23	.21	6.44
	$\psi(\hat{\theta}_4)$	.85	$.99 \times 10^{-2}$	.19	.33	7.55
40	$\psi(\hat{\theta}_1)$	.85	3.95×10^{-2}	.28	−.20	1
	$\psi(\hat{\theta}_2)$	.85	$.82 \times 10^{-2}$	.30	.11	4.82
	$\psi(\hat{\theta}_3)$	.85	$.59 \times 10^{-2}$	.24	.27	6.69
	$\psi(\hat{\theta}_4)$	.85	$.53 \times 10^{-2}$	.26	.01	7.45
60	$\psi(\hat{\theta}_1)$	.84	2.56×10^{-2}	.08	−.01	1
	$\psi(\hat{\theta}_2)$	.85	$.51 \times 10^{-2}$	.15	.17	5.02
	$\psi(\hat{\theta}_3)$	.85	$.36 \times 10^{-2}$	.08	.07	7.11
	$\psi(\hat{\theta}_4)$	.85	$.30 \times 10^{-2}$	.02	−.02	8.53

TABLE 3 At-the-Money Call Option Simulation Results

Sample Size	Statistic	Mean	Variance	Skew	Kurtosis	Sample Relative Efficiency
5	$\psi(\hat{\theta}_1)$	2.87	57.24×10^{-2}	.32	−.07	1
	$\psi(\hat{\theta}_2)$	2.97	10.76×10^{-2}	.39	.12	5.32
	$\psi(\hat{\theta}_3)$	2.99	7.30×10^{-2}	.29	.08	7.84
	$\psi(\hat{\theta}_4)$	3.01	6.86×10^{-2}	.30	.20	8.34
10	$\psi(\hat{\theta}_1)$	2.97	29.12×10^{-2}	.32	.38	1
	$\psi(\hat{\theta}_2)$	2.99	6.24×10^{-2}	.36	.25	4.66
	$\psi(\hat{\theta}_3)$	2.99	4.10×10^{-2}	.16	−.20	7.10
	$\psi(\hat{\theta}_4)$	3.00	3.72×10^{-2}	.25	.22	7.83
20	$\psi(\hat{\theta}_1)$	2.99	13.71×10^{-2}	.10	−.04	1
	$\psi(\hat{\theta}_2)$	3.00	2.95×10^{-2}	.16	.39	4.65
	$\psi(\hat{\theta}_3)$	3.00	2.07×10^{-2}	.15	.14	6.62
	$\psi(\hat{\theta}_4)$	3.01	1.79×10^{-2}	.11	.25	7.66
40	$\psi(\hat{\theta}_1)$	3.01	7.10×10^{-2}	.15	−.22	1
	$\psi(\hat{\theta}_2)$	3.01	1.47×10^{-2}	.24	.05	4.83
	$\psi(\hat{\theta}_3)$	3.01	1.06×10^{-2}	.18	.22	6.70
	$\psi(\hat{\theta}_4)$	3.01	$.94 \times 10^{-2}$	.21	−.02	7.55
60	$\psi(\hat{\theta}_1)$	3.00	4.64×10^{-2}	−.04	.00	1
	$\psi(\hat{\theta}_2)$	3.01	$.91 \times 10^{-2}$	.10	.11	5.10
	$\psi(\hat{\theta}_3)$	3.01	$.65 \times 10^{-2}$	.04	.04	7.14
	$\psi(\hat{\theta}_4)$	3.01	$.54 \times 10^{-2}$	−.02	−.03	8.59

TABLE 4 In-the-Money Call Option Simulation Results

Sample Size	Statistic	Mean	Variance	Skew	Kurtosis	Sample Relative Efficiency
5	$\phi(\hat\theta_1)$	6.62	25.31 × 10 [illegible]	1.04	1.13	1
	$\phi(\hat\theta_2)$	6.63	5.44 × 10 [illegible]	.70	.53	4.65
	$\phi(\hat\theta_3)$	6.64	3.69 × 10 [illegible]	.55	.31	6.86
	$\phi(\hat\theta_4)$	6.65	3.53 × 10 [illegible]	.56	.47	7.17
10	$\phi(\hat\theta_1)$	6.65	14.26 × 10 [illegible]	.89	1.55	1
	$\phi(\hat\theta_2)$	6.64	3.18 × 10 [illegible]	.61	.65	4.48
	$\phi(\hat\theta_3)$	6.65	2.06 × 10 [illegible]	.34	−.15	6.92
	$\phi(\hat\theta_4)$	6.64	1.89 × 10 [illegible]	.45	.38	7.54
20	$\phi(\hat\theta_1)$	6.65	6.74 × 10 [illegible]	.45	.21	1
	$\phi(\hat\theta_2)$	6.64	1.50 × 10 [illegible]	.35	.69	4.49
	$\phi(\hat\theta_3)$	6.64	1.05 × 10 [illegible]	.30	.30	6.42
	$\phi(\hat\theta_4)$	6.64	.91 × 10 [illegible]	.25	.42	7.41
40	$\phi(\hat\theta_1)$	6.65	3.60 × 10 [illegible]	.38	−.14	1
	$\phi(\hat\theta_2)$	6.65	.75 × 10 [illegible]	.35	.17	4.80
	$\phi(\hat\theta_3)$	6.64	.54 × 10 [illegible]	.29	.31	6.66
	$\phi(\hat\theta_4)$	6.64	.48 × 10 [illegible]	.30	.04	7.50
60	$\phi(\hat\theta_1)$	6.64	2.31 × 10 [illegible]	.17	.03	1
	$\phi(\hat\theta_2)$	6.64	.46 × 10 [illegible]	.20	.23	5.02
	$\phi(\hat\theta_3)$	6.64	.33 × 10 [illegible]	.12	.09	7.00
	$\phi(\hat\theta_4)$	6.64	.27 × 10 [illegible]	.05	−.01	8.55

In many cases we will want more than a point estimate of the Black-Scholes call option price. An approximate $(1-\alpha) \times 100\%$ confidence interval for $\phi(\theta)$ is simply

$$\phi(\hat\theta) \pm Z_{\alpha/2}\sqrt{-\left(\frac{d^2 \ln L}{d\theta^2}\right)^{-1}_{\theta=\hat\theta}},$$

where $\hat\theta$ is the maximum likelihood estimate and $Z_{\alpha/2}$ is the $(1-\alpha/2) \times 100$th percentile of the standard normal distribution. Since the maximum likelihood call option price estimator possesses minimal bias and maximum statistical efficiency, the resulting confidence intervals will be precise and symmetric. Wilks (1938) has formally established the optimality of confidence intervals derived from maximum likelihood estimation. These intervals are symmetric and for large samples are shortest, on average.

This methodology should be contrasted with the previous analysis of Boyle and Ananthanarayanan (1977). Comparatively, their call option price confidence intervals are less precise and nonsymmetric. Recall that Boyle and Ananthanarayanan's analysis was based on the observation that from a sample of n independent observations on closing prices

$$(n-1)\frac{s^2}{\theta} \sim \chi^2_{(n-1)},$$

where

$$s^2 = \sum_{i=1}^{n} \frac{(x_i - \bar{x})^2}{n - 1}.$$

A $(1 - \alpha) \times 100\%$ confidence interval for θ follows:

$$\frac{(n - 1)s^2}{\chi^2_{(n-1),1-\alpha/2}} \leq \theta \leq \frac{(n - 1)s^2}{\chi^2_{(n-1),\alpha/2}},$$

where $\chi^2_{(n-1),1-\alpha/2}$ and $\chi^2_{(n-1),\alpha/2}$ are, respectively, the $(1 - \alpha/2) \times 100$th and the $(\alpha/2) \times 100$th percentiles of the χ^2 distribution with $(n - 1)$ degrees of freedom. The nonlinear Black-Scholes call option price transformation is directly applied to the preceding confidence interval to give

$$\psi\left[\frac{(n - 1)s^2}{\chi^2_{(n-1),1-\alpha/2}}\right] \leq \psi(\theta) \leq \psi\left[\frac{(n - 1)s^2}{\chi^2_{(n-1),\alpha/2}}\right],$$

a $(1 - \alpha) \times 100\%$ confidence interval for $\psi(\theta)$, the Black-Scholes call option price. However, this methodology is imprecise in a number of respects. For all sample sizes the χ^2 distribution is nonsymmetric, implying that a nonsymmetric θ confidence interval will result. The $\psi(\theta)$ confidence interval nonsymmetry is further exaggerated, since a nonlinear transformation is directly applied to an already nonsymmetric θ confidence interval. Finally, the Boyle and Ananthanarayanan analysis is based solely on closing prices. By incorporating further information, in particular daily high and low security prices, we can construct more precise call option price confidence intervals.

While the maximum likelihood methodology yields optimal confidence intervals for large sample sizes, the properties of the corresponding confidence intervals derived from moderate-sized samples remains to be explored. From a practitioner's perspective this is the more relevant basis of comparing Boyle and Ananthanarayanan's methodology with the maximum likelihood methodology. To that end, 95% confidence intervals for the previously hypothesized at-the-money call option with true price of \$3.01 will be constructed under the alternative methodologies, given 20, 40, and 60 days of simulated data. The Boyle and Ananthanarayanan confidence intervals are based on the assumption that s^2 equals the true value of θ. This assumption simply follows the convention of their analysis. Evoking asymptotic theory, the maximum likelihood confidence intervals make use of the estimate of the standard deviation of $\psi(\hat{\theta}_4)$. The results are presented in table 5. For all sample sizes, the maximum likelihood methodology provides a far more precise and symmetric and indeed a far more powerful 95% confidence interval.

TABLE 5 Comparison of Confidence Intervals ($)

Sample Size	Boyle and Ananthanarayanan Confidence Intervals	Maximum Likelihood Confidence Intervals
20	2.43–4.14	2.75–3.27
	(1.71)	(.52)
40	2.57–3.71	2.82–3.20
	(1.14)	(.38)
60	2.64–3.56	2.87–3.15
	(.92)	(.28)

NOTE.—Length of confidence intervals in parentheses.

V. Conclusions

Assuming security price dynamics are governed by a diffusion process and given the publicly most available form of data, this paper has derived the maximum likelihood estimator of security price volatility. Controlled simulation studies established that, for reasonable sample sizes, this procedure yields essentially unbiased estimates with the highest degree of statistical efficiency. Subsequently, the corresponding maximum likelihood estimator of the Black-Scholes call option price was formulated. Controlled simulation studies indicated the preferability of this estimation procedure. Moreover, the asymptotic properties of the maximum likelihood estimator allowed the construction of the shortest and, as such, the optimal confidence intervals for the Black-Scholes call option price.

A number of caveats must yet be raised regarding the practical limitations of this and all other proposed high-low estimators of security price volatility. Without a doubt, their usefulness depends critically on actual security price dynamics being governed by the posited diffusion process. From a practical perspective, questions exist whether observed security price highs and lows correspond to actual security price highs and lows. Finally, security price volatility estimation procedures must more fully integrate the closed market effect. Answers to these queries are beyond the scope of the present analysis. Further research in these directions is clearly required.

What should be emphasized, however, is that these criticisms apply to all estimation procedures that fall within the class of high-low security volatility estimators. These criticisms aside, given the validity of the posited security price process and the reliability of the available data, this paper has established the advantages of the maximum likelihood security price volatility estimator.

Appendix A

Joint Density Function

The joint density of the daily closing, high, and low security prices conditional on $\theta \equiv \sigma^2$ is given by

$$f(x, a, b; \theta) = \sum_{n=1}^{\infty} \exp[-\delta(n)\theta][\alpha(n)\theta^2 + \beta(n)\theta + \gamma(n)],$$

where

$$\alpha(n) = \frac{2n^4\pi^4}{(a+b)^7} \sin\left[\frac{n\pi\,(b+x)}{a+b}\right] \sin\left(\frac{n\pi\, b}{a+b}\right),$$

$$\beta(n) = \frac{2n^3\pi^3(a-b)}{(a+b)^6} \sin\left[\frac{n\pi\,(2b+x)}{a+b}\right]$$

$$- \frac{4n^3\pi^3}{(a+b)^6} \sin\left(\frac{n\pi\, b}{a+b}\right) \cos\left[\frac{n\pi\,(b+x)}{a+b}\right]$$

$$- \frac{10\, n^2\pi^2}{(a+b)^5} \sin\left(\frac{n\pi\, b}{a+b}\right) \sin\left[\frac{n\pi\,(b+x)}{a+b}\right],$$

$$\gamma(n) = \frac{2n^2\pi^2}{(a+b)^5} (bx - 2ab - ax) \cos\left[\frac{n\pi\,(2b+x)}{a+b}\right]$$

$$+ \frac{4n\pi\,(b-a)}{(a+b)^4} \sin\left[\frac{n\pi\,(2b+x)}{a+b}\right]$$

$$+ \frac{8n\pi\, x}{(a+b)^4} \sin\left(\frac{n\pi\, b}{a+b}\right) \cos\left[\frac{n\pi\,(b+x)}{a+b}\right]$$

$$+ \left[\frac{4}{(a+b)^3} - \frac{2n^2\pi^2x^2}{(a+b)^5}\right] \sin\left(\frac{n\pi\, b}{a+b}\right) \sin\left[\frac{n\pi\,(b+x)}{a+b}\right].$$

and

$$\delta(n) = \frac{n^2\pi^2}{2(a+b)^2}.$$

Appendix B

Implementation of the Maximum Likelihood Estimation Procedure

As previously stated, the maximum likelihood estimator, $\hat{\theta}$, is that value of θ which maximizes

$$\ln L(\theta) = \sum_{i=1}^{M} \ln \sum_{n=1}^{\infty} [\alpha_i(n)\theta^2 + \beta_i(n)\theta + \gamma_i(n)] \exp[-\delta_i(n)\theta]. \qquad \text{(B1)}$$

First consider the numerical calculation of this infinite sum. Rather than examining the entire expression, we restrict our attention to the transition density,

$$p(x, t) = \sum_{n=1}^{\infty} \frac{2}{a+b} \sin \frac{n\pi b}{a+b} \sin \frac{n\pi (b+x)}{a+b} \exp\left[\frac{-n^2\pi^2\theta}{2(a+b)^2} t\right]. \quad \text{(B2)}$$

Clearly, adequate approximation of $p(x, t)$ will imply adequate approximation of $\ln L(\theta)$. The infinite sum in (B2) shall be truncated at N and the resultant approximation error denoted by $G(N)$. Since the sine function is bounded in absolute value by unity, it follows that

$$|G(N)| \leq \sum_{n=N+1}^{\infty} \frac{2}{a+b} \cdot \exp\left[-\frac{n^2\pi^2\theta}{2(a+b)^2}\right].$$

However, the function $2/(a+b) \cdot \exp[-n^2\pi^2\theta/2(a+b)^2]$ is monotonically decreasing in n, implying that we may bound this sum by the corresponding integral to reveal

$$|G(N)| \leq \int_{x=N}^{\infty} \frac{2}{(a+b)} \cdot \exp\left[-\frac{x^2\pi^2\theta}{2(a+b)^2}\right] dx.$$

Furthermore, by the change of variable

$$w = \frac{x\pi\theta^{1/2}}{(a+b)},$$

we have

$$|G(N)| \leq \int_{w=\frac{N\pi\theta^{1/2}}{(a+b)}}^{\infty} 2\exp\left[-\frac{w^2}{2}\right] dw.$$

Using the following well-known inequality (Abromowitz and Stegun 1968),

$$\int_{x}^{\infty} \exp\left[-\frac{w^2}{2}\right] dw < \frac{\exp\left[-\frac{x^2}{2}\right]}{x} \quad \text{for } x > 0,$$

it follows that

$$|G(N)| < \frac{2(a+b)}{N\pi^2\theta} \exp\left[-\frac{N^2\pi^2\theta}{2(a+b)^2}\right].$$

We now have a bound on the truncation error. With this error bound as a framework, a series of experiments were performed to select an optimal truncation point. For the purposes of the current work, truncation at $N = 7$ provided double precision accuracy. The maximum likelihood procedure has been reduced to maximizing:

$$\ln L_N(\theta) = \sum_{i=1}^{M} \ln \sum_{n=1}^{N} \exp[-\delta_i(n)\theta]\,[\alpha_i(n)\theta^2 + \beta_i(n)\theta + \gamma_i(n)].$$

A necessary condition for a maximum is that $d \ln L_N(\theta)/d\theta = 0$. The solution of this nonlinear equation may be found using a Newton-Raphson iterative procedure. This technique linearizes $d \ln L_N(\theta)/d\theta$ as a function of θ according to its Taylor series expansion. A sequence $\{\theta_0, \theta_1, \theta_2, \ldots,\}$ that converges to $\hat{\theta}$ is formed as follows:

$$\theta_{i+1} = \theta_i - \frac{d \ln L_N(\theta_i)}{d\theta} \Bigg/ \frac{d^2 \ln L_N(\theta_i)}{d\theta^2}.$$

It is important to note that the first and second derivatives of $\ln L_N(\theta)$ are simple functions of θ and $\ln L_N(\theta)$. As such, they can be conveniently computed simultaneously with $\ln L_N(\theta)$ to allow an efficient computer algorithm. The sequence is terminated when a solution is found to within predesignated error limits. For every case in the simulations, $d^2 \ln L_N(\theta)/d\theta^2$ was observed and found to be negative, indicating a local maximum of $\ln L_N(\theta)$. The choice of a starting point in this analysis is quite important. When the Garman-Klass estimator was employed, convergence was always achieved within four iterations. Experiments were performed for a wide range of starting values, with the result that the algorithm converged to the same solution in a small number of iterations. Based on experimental evidence, the likelihood function is everywhere concave downward. Consequently, there is a unique maximum which is always obtained using this numerical procedure.

Appendix C

Simulation of a Wiener Process

Let $\{X(t): t \geq 0\}$ denote a Wiener process with zero drift and diffusion coefficient σ. Our object is to simulate $X(t)$ for $0 \leq t \leq 1$ and then to note the high, low, and close for this process: $\max_{0 \leq t \leq 1} X(t)$, $\min_{0 \leq t \leq 1} X(t)$, and $X(1)$, respectively.

For convenience, denote one time unit as a "day." We split the day into k intervals and, by simulating a Brownian increment for each interval, we obtain the following approximation for $X(t)$:

$$\xi(t) = \sum_{j=1}^{[kt]} \eta_j, \; 0 \leq t \leq 1.$$

The notation $[\cdot]$ designates integer part and $\{\eta_j\}$ is a simulated sequence of independent identically distributed (i.i.d.) uniform random variables with mean zero and variance σ^2/k. By appealing to the central limit theorem and according to Shreider (1967), $\xi(t)$ converges in quadratic mean to $X(t)$ for $0 \leq t \leq 1$ as $k \to \infty$.

As a preliminary to generating $\{\eta_j\}$, a sequence $\{\mu_j\}$ of i.i.d. pseudorandom variables with distribution uniform on $[0, 1]$ is formed. The FORTRAN subroutine URAND was employed for this purpose. This random number generator uses a linear congruential method whose integer sequence is given by $u_{j+1} = 834{,}314{,}861\ \mu_j + 453{,}816{,}693 \pmod{2^{31}}$. For details, see Forsythe, Malcolm, and Moler (1971). The transformation from μ_j to η_j is given by $\eta_j = (\sigma/\sqrt{k})(\mu_j - .5)$, for each j.

To simulate a Wiener process adequately, the number of intervals k must be rather large. Although it is straightforward to simulate the close $X(1)$, the complex probablistic structure between the high and the low makes simulation difficult. We decided to choose $k = 100{,}000$, a sufficiently large number of intervals. Unfortunately this selection creates two difficulties: the computer expense in numerous calls of GRAND and the potential for serial correlation in the random number generator. A sampling design was adopted to counter both these problems.

Each day is split into 100 periods and each period is further split into 1,000 intervals. To compute the high, low, and close for a particular period requires 1,000 calls of the subroutine GRAND. A data base of high, low, and close for 10,000 periods was created and stored. To generate a single day's simulated data, the subroutine URAND was used to select at random 100 periods from the data base. However, Brownian motion increments are independent; as a result, the 100 periods were pieced together to form the high, low, and close for one day.

Clearly, any serial correlation between pairs of periods will be diffused, and furthermore, this procedure of random sampling in blocks reduces computer expenses by a factor of 100. Moreover, to ascertain the quality of this experimental design, further simulation studies were performed and revealed that sample highs and lows corresponded (apart from sampling error) to the population highs and lows under the diffusion model. In addition, the distribution of the sample close is also statistically indistinguishable from its population counterpart. On the basis of these tests and the results of our study we conclude that our simulation procedures provide sample data consistent with the posited diffusion process.

References

Abromowitz, M., and Stegun, I. A. 1968. *Handbook of Mathematical Functions*. Washington, D.C.: National Bureau of Standards.

Black, F., and Scholes, M. 1973. The pricing of options and corporate liabilities. *Journal of Political Economy* 81:637–54.

Boyle, P. P., and Ananthanarayanan, A. L. 1977. The impact of variance estimation in option valuation models. *Journal of Financial Economics* 5:375–87.

Cox, D. R., and Miller, H. D. 1972. *The Theory of Stochastic Processes*. London: Chapman & Hall.

Cramer, H. 1946. *Mathematical Methods of Statistics*. Princeton, N.J.: Princeton University Press.

Forsythe, G. E.; Malcolm, M. A.; and Moler, C. B. 1977. *Computer Methods for Mathematical Computations*. Englewood Cliffs, N.J.: Prentice-Hall.

Garman, M. B., and Klass, M. J. 1980. On the estimation of security price volatilities from historical data. *Journal of Business* 53, no. 1: 67–78.

Merton, R. C. 1980. On estimating the expected return on the market. *Journal of Financial Economics* 8:323–61.

Parkinson, M. 1980. The extreme value method for estimating the variance of the rate of return. *Journal of Business* 53, no. 1: 61–65.

Rosenberg, B. 1972. The behavior of random variables with non stationary variance and the distribution of security prices. Working Paper no. 11. Berkeley: University of California, December.

Shreider, Y. A., ed. 1967. *The Monte Carlo Method*. London: Pergamon.

Wilks, S. S. 1938. Shortest average confidence intervals from large samples. *Annals of Mathematical Statistics* 9:166–75.