# Stock Price Distributions with Stochastic Volatility: An Analytic Approach

**Elias M. Stein**
Princeton University

**Jeremy C. Stein**
Massachusetts Institute of Technology

***We study the stock price distributions that arise when prices follow a diffusion process with a stochastically varying volatility parameter. We use analytic techniques to derive an explicit closed-form solution for the case where volatility is driven by an arithmetic Ornstein–Uhlenbeck (or AR1) process. We then apply our results to two related problems in the finance literature: (i) options pricing in a world of stochastic volatility, and (ii) the relationship between stochastic volatility and the nature of "fat tails" in stock price distributions.***

In this article, we study the stock price distributions that arise when prices follow a diffusion process with a stochastically varying volatility parameter, as described in the following two equations:

$$dP = \mu P\,dt + \sigma P\,dz_1 \tag{1}$$

and

$$d\sigma = -\delta(\sigma - \theta)\ dt + k\,dz_2, \tag{2}$$

where $P$ is the stock price, $\sigma$ is the "volatility" of the stock, $k$, $\mu$, $\delta$, and $\theta$ are fixed constants, and $dz_1$ and

We thank Ron Henderson for invaluable research assistance. We also thank Michael Brennan, Peter Freund, Ken Froot, Bruce Petersen, Dan Stroock, and an anonymous referee for helpful comments and discussions. Address correspondence to Jeremy C. Stein, Sloan School of Management, MIT, Cambridge, MA 02139.

*The Review of Financial Studies* 1991 Volume 4, number 4, pp. 727–752

$dz_2$ are two independent Wiener processes. Thus, the model is one where volatility is governed by an arithmetic Ornstein–Uhlenbeck (or AR1) process, with a tendency to revert back to a long-run average level of $\theta$. We use analytic techniques (related to the heat equation for the Heisenberg group) to derive a closed-form solution for the distribution of stock prices in this case.

Our primary interest in doing so is to generate an options pricing formula that is appropriate for the case where volatility follows an autoregressive stochastic process. A large and growing literature suggests that this case is empirically relevant. Although the empirical literature offers many different models for time-varying volatility, of which the AR1 is but one example, the AR1 model provides a natural starting point for the types of questions we address here.[1] It is parsimonious enough that the analysis is tractable, yet (as we argue in more detail below) it captures many of the documented features of volatility data.

Recently, interesting papers by Johnson and Shanno (1987), Wiggins (1987), and Hull and White (1987) have also examined options pricing in a world where stock price dynamics are similar to those given by Equations (1) and (2). The first two papers use numerical methods to determine options prices. In the third, Hull and White solve explicitly for the options price by using a Taylor expansion about $k = 0$ (i.e., about the point where volatility is nonstochastic). It is not clear that such an expansion provides a good approximation to options prices when $k$ is significantly greater than zero. Furthermore, Hull and White only apply this series solution for the case $\delta = 0$. (They use numerical methods to study the case of nonzero $\delta$.)

Although our closed-form solution is quite cumbersome, it is composed entirely of elementary mathematical functions. Consequently, it is readily and directly applied on a desktop computer. We are thus able to avoid using more burdensome numerical methods, or any assumptions about $k$ being close to zero. Also, our method is capable of handling a nonzero mean reversion parameter $\delta$, which should be valuable, given the empirical evidence that volatility is strongly mean-reverting.

In addition to deriving an exact closed-form solution for the stock price distribution, we also use analytic techniques to develop an approximation to the distribution. The approximation has the advantage of being even less computationally demanding than the exact

[1] Empirical articles that model volatility as an AR1 process include Poterba and Summers (1986), Stein (1989), and Merville and Pieptea (1989). (The latter allows the AR1 process to be displaced by white noise.) Alternative models include the ARCH model of Engle (1982) and its descendants such as Bollerslev (1986), Engle and Bollerslev (1986), and Engle, Lilien, and Robins (1987). [See also French, Schwert, and Stambaugh (1987) and Schwert (1987, 1989).]

solution. For most parameter values, we also find that the approximation is reasonably accurate—it leads to options prices close to those obtained with the exact formula.

A by-product of our analysis is that it allows us to draw a direct link between the parameters of the volatility process and the extent to which stock price distributions have "fat tails" compared to the lognormal distribution [see, e.g., Fama (1963, 1965) and Mandlebrot (1963)]. We are able to show explicitly how the shape of stock price distributions depends on the parameters of our Equation (2), thereby tracing fat tails back to their "primitive origins," the constants $\delta$, $\theta$, and $k$.[2]

The remainder of the article is organized as follows. Our principal results for exact and approximate stock price distributions are presented in Section 1. In Section 2, these results are applied to options pricing. Using the empirical literature on stochastic volatility as a guide, we select a range of "reasonable" parameter values and compare the prices generated by our model to Black–Scholes (1973) prices. In Section 3, we explore the connection between our model's parameters and the degree to which stock price distributions over different time horizons have fat tails. In Section 4, we conclude and discuss some possible extensions of our work.

The problems studied in this article require a substantial amount of mathematical analysis. In order not to interfere with the presentation of the main ideas, most of it is omitted from the text. The Appendixes contain a brief review of the important derivations. Further detail is available from the authors on request.

Before proceeding, we ought to comment on our assumption that volatility is driven by an arithmetic process, which raises the possibility that $\sigma$ can become negative. This formulation is equivalent to putting a reflecting barrier at $\sigma = 0$ in the volatility process, since $\sigma$ enters everywhere else in squared fashion. Although this is not a very natural feature, geometric models for $\sigma$ are much less analytically tractable.[3] Furthermore, as we argue in Section 2.1 below, any objections to the arithmetic process are more theoretical than practical. For a wide range of relevant parameter values, the probability of actually reaching the point $\sigma = 0$ is so small as to be of no significant consequence.

[2] There are analogous results in the literature for different models of volatility. For example, Praetz (1972) and Blattberg and Gonedes (1974) show that if $\sigma^2$ follows an inverted gamma distribution, then prices are distributed as a log $t$ [see also Clark (1973)]. Engle (1982) computes the kurtosis of an ARCH process as an explicit function of the ARCH parameters.

[3] While we have been able to solve a model where $\sigma$ follows a geometric random walk (and the results are considerably more computationally cumbersome than those reported here), we have been unable to make any progress on the case where $\sigma$ follows a geometric Ornstein–Uhlenbeck process. Since mean reversion appears to be one of the most important empirical characteristics of volatility, the practical usefulness of the geometric random walk model is unclear.

# 1. Exact and Approximate Stock Price Distributions

## 1.1 The closed-form exact solution

An explicit, closed-form solution for the stock price distribution corresponding to Equations (1) and (2) is presented below. As a normalization, we set the initial stock price, $P_0$, equal to unity.

We start by defining the following new variables:

$$A \equiv -\delta/k^2, \qquad B = \theta\delta/k^2, \qquad C = -\lambda/k^2 t. \tag{3}$$

$A$ and $B$ are simply functions of our primitive parameters; $C$ also contains $\lambda$, which is a dummy variable to be used in defining an integral expression below.

Next, we define

$$a \equiv (A^2 - 2C)^{1/2}, \qquad b \equiv -A/a, \tag{4}$$

$$L \equiv -A - a\left(\frac{\sinh(ak^2t) + b\cosh(ak^2t)}{\cosh(ak^2t) + b\sinh(ak^2t)}\right), \tag{5}$$

$$M \equiv B\left\{\frac{b\sinh(ak^2t) + b^2\cosh(ak^2t) + 1 - b^2}{\cosh(ak^2t) + b\sinh(ak^2t)} - 1\right\}, \tag{6}$$

$$\begin{aligned} N \equiv{} & \frac{a-A}{2a^2}[a^2 - AB^2 - B^2a]k^2t + \frac{B^2[A^2 - a^2]}{2a^3} \\ & \times\left\{\frac{(2A + a) + (2A - a)e^{2ak^2t}}{(A + a + (a - A)e^{2ak^2t})}\right\} \\ & + \frac{2AB^2[a^2 - A^2]e^{ak^2t}}{a^3\left(A + a + (a - A)e^{2ak^2t}\right)} \\ & - \frac{1}{2}\log\left\{\frac{1}{2}\left(\frac{A}{a} + 1\right) + \frac{1}{2}\left(1 - \frac{A}{a}\right)e^{2ak^2t}\right\}, \end{aligned} \tag{7}$$

and

$$I = \exp(L\sigma_0^2/2 + M\sigma_0 + N). \tag{8}$$

Notice that $I$, in addition to being a function of the primitive parameters $\delta$, $\theta$, $k$, $t$, and $\sigma_0$, also depends on the dummy variable $\lambda$. We denote this relationship by writing $I$ as $I(\lambda)$. We now write down an expression for the stock price distribution in two steps. First, we define $S_0(P, t)$ as the time $t$ stock price distribution in the special case where the stock price drift $\mu = 0$. It is given by

$$S_0(P, t) = (2\pi)^{-1} P^{-3/2} \times \int_{\eta=-\infty}^{\infty} I\left(\left(\eta^2 + \frac{1}{4}\right)\frac{t}{2}\right) e^{i\eta \log P}\, d\eta. \tag{9}$$

The time $t$ stock price distribution for the more general case of a nonzero $\mu$, which we denote simply by $S(P, t)$, is then

$$S(P, t) = e^{-\mu t} S_0(Pe^{-\mu t}). \tag{10}$$

Although the formula for $S(P, t)$ is complicated, it is composed entirely of elementary functions and requires only a single integration. Also, note that $S(P, t)$ is a conditional distribution, conditional upon the current stock price and current volatility—more comprehensive notation would involve writing this distribution as $S(P, t \mid P_0, \sigma_0)$.

### 1.2 An approximate solution

In some cases, it may be useful to have a somewhat simpler approximation to $S(P, t)$. In order to derive such an approximation, we note that our exact distribution can always be expressed as an average of lognormal distributions, averaged via a mixing distribution. (This is proved in Appendix A.) That is,

$$S(P, t) = \int L(\sigma)\, m_t(\sigma)\, d\sigma, \tag{11}$$

where $L(\sigma)$ is a lognormal with the same mean as $S(P, t)$ and volatility $\sigma$, and $m_t(\sigma)$ is a mixing distribution. The $t$ subscript on $m_t(\sigma)$ emphasizes the fact that the mixing distribution depends on the time horizon.

Our approximation technique works by approximating this mixing distribution $m_t(\sigma)$. In Appendix C, we show that $m_t(\sigma)$ can be well approximated (in a sense we make precise) by $\hat{m}_t(\sigma)$, which has the simple form

$$\hat{m}_t(\sigma) = \rho e^{-\alpha\sigma^2} e^{-\beta/\sigma^2}, \tag{12}$$

where the parameters $\rho$, $\alpha$, and $\beta$ are defined in Appendix C. With $\hat{m}_t(\sigma)$ in hand, our approximate stock price distribution $\hat{S}(P, t)$ is given by

$$\hat{S}(P, t) = \int L(\sigma)\, \hat{m}_t(\sigma)\, d\sigma. \tag{13}$$

## 2. Application to Options Pricing

In this section, we apply our results to the pricing of European stock options. It can be shown that, given our assumptions, the price of an option $F$ must satisfy the partial differential equation

$$\tfrac{1}{2}\sigma^2 P^2 F_{pp} + rPF_p - rF + F_t + \tfrac{1}{2}k^2 F_{\sigma\sigma} + F_\sigma[-\delta(\sigma - \theta) - \phi k] = 0. \quad (14)$$

Here $\phi$ denotes the market price of the stock's volatility risk and $r$ is the riskless interest rate. The presence of $\phi$ in Equation (14) reflects the fact that with stochastic volatility, one cannot use arbitrage arguments to eliminate investor risk preferences from the options pricing problem, because the volatility $\sigma$ is itself not a traded asset.[4]

As we demonstrate below, our analytic results allow us to solve (14) both in the case where $\phi$ is zero, as well as in the case where $\phi$ is a nonzero constant. Before doing so, however, it is worth noting when either of these two assumptions can be justified in the context of a specific equilibrium model. Wiggins (1987) contains a thorough discussion of the equilibrium determinants of $\phi$, drawing on the results of Cox, Ingersoll, and Ross (1985). Wiggins points out that, in general, $\phi$ need not be constant, and indeed may not be expressible in a closed form. However, he goes on to identify some interesting special cases that obtain when there is a representative consumer with log utility.

With log utility, $\phi = 0$ if the option in question is an option on the market portfolio. Somewhat more generally, Wiggins argues that log utility also leads to a constant (though possibly nonzero) $\phi$ for individual stocks, so long as the following underlying moments themselves remain constant: the standard deviation of the market portfolio, and the pairwise correlation coefficients between the individual stock's returns, its volatility, and the return on the market portfolio. These results imply that the assumption of a constant $\phi$ can indeed be compatible with a fully specified (but somewhat restrictive) equilibrium model. In the case where $\phi = 0$, the pricing equation simplifies to

$$\tfrac{1}{2}\sigma^2 P^2 F_{pp} + rPF_p - rF + F_t + \tfrac{1}{2}k^2 F_{\sigma\sigma} + F_\sigma[-\delta(\sigma - \theta)] = 0. \quad (14')$$

Equation (14′) does not depend on risk preferences. Thus, in a world where $\phi = 0$, we can calculate the option price by assuming that risk neutrality prevails. This implies that the price of a European call is given by

$$F_0 = e^{-rt} \int_{P=K}^{\infty} [P - K]\, S(P, t \mid \delta, r, k, \theta)\, dP. \quad (15)$$

[4] Very similar expressions appear in Wiggins (1987) and Hull and White (1987). In particular, our PDE is a simplification of Wiggins' equation (8) (p. 355) to the case where $dz_1$ and $dz_2$ are uncorrelated and where $\sigma$ follows an arithmetic, rather than geometric, process. Our model is not as general as Wiggins' because we have been unable to apply our analytic techniques when $dz_1$ and $dz_2$ are correlated. However, it may be possible to capture the tendency for volatility and stock prices to move together even without assuming an instantaneous correlation between $dz_1$ and $dz_2$. This might be accomplished in our framework by using a constant-elasticity-of-volatility generalization of Equation (1). We discuss this briefly in Section 4.

The subscript 0 on $F$ emphasizes the fact that Equation (15) applies only to the case where $\phi = 0$. The stock price distribution $S(P, t)$ in this equation is generated using the parameters $\delta$, $k$, and $\theta$, along with the assumption that the stock's drift equals the riskless rate $r$. Thus, the mean of $S(P, t)$ in (15) is equal to $P_0 e^{rt}$.[5]

Now consider the somewhat more general case of a nonzero but constant $\phi$. From a comparison of (14) and (14′), it is apparent that if the solution to (14′) is given by (15), then the solution to (14) must be given by

$$F = e^{-rt} \int_{P=K}^{\infty} [P - K]\, S(P, t \mid \delta, r, k, \hat{\theta}), \tag{16}$$

where $\hat{\theta} = \theta - \phi k/\delta$. Therefore, in addition to setting the stock price drift to $r$, we also modify the parameter $\theta$ to account for the effect of volatility risk on options prices. When the volatility risk premium $\phi$ is positive, $\hat{\theta}$ is lower than $\theta$, and options prices are lower, all else being equal.

Before presenting some sample options prices, we briefly discuss two practical issues. The first concerns the calibration of the model—that is, the choice of appropriate parameter values. The second concerns the implementation of the model on a computer.

## 2.1 Calibrating the model

In order to choose reasonable values for the parameters $\theta$, $k$, and $\delta$, we draw on the existing empirical literature on the time-series properties of volatility, focusing upon those articles that use an AR1 model for volatility. While an AR1 model cannot be expected to fit the data as well as a less parsimonious one, there is some evidence that suggests it does quite a good job, at least for aggregate stock indices. For example, Stein (1989) finds that when modeling the implied volatility of S&P 100 index options, the corrected $R^2$ from an AR1 model is the *same* as that from a model that includes eight lags. Using a different data set, Poterba and Summers (1986) also come to the conclusion that an AR1 model provides a good description of the time-series behavior of volatility.[6]

A related concern is the appropriateness of using an arithmetic versus geometric specification for the volatility process. However,

---

[5] Hull and White produce a similar result [equation (6), p. 283] also by effectively assuming that $\phi = 0$.

[6] Using implied volatilities for individual stocks, Merville and Pieptea (1989) argue that a better fit is obtained by allowing the AR1 process to be displaced by white noise. One plausible interpretation is that "true" volatility follows an AR1, but that their implied volatilities contain significant white noise measurement errors.

note that Stein (1989) finds no evidence of skewness in the implied volatilities of S&P 100 index options, concluding that it is reasonable to model the volatility process in levels. In addition, empirically reasonable parameter values imply a very small probability that an arithmetic volatility process will ever reach zero.

Two studies are used as principal sources for parameter values, those of Stein (1989) and Merville and Pieptea (1989). Both studies use options-implied volatilities as their data, rather than relying on volatilities estimated from stock price returns. The former focuses on S&P 100 index options, while the latter looks primarily at 25 individual stocks.[7]

Using data from 1983 to 1987, Stein (1989) finds that index volatility averages between about 15 and 20 percent, depending on the sample period. The half-life of a volatility shock is approximately one month for the entire sample, corresponding to a $\delta$ of 8. However, in some subsamples, the half-life is as short as two weeks, corresponding to a $\delta$ of 16.[8] Estimates of $k$ range from 0.15 to 0.30.

The individual stocks examined by Merville and Pieptea (1989) have (not surprisingly) higher average implied volatilities than the index, generally ranging between about 25 and 35 percent. Mean reversion is typically stronger than with the index, with a median value of $\delta$ in the range of 14, and several individual observations over 20. The parameter $k$ also tends to be higher, often exceeding 0.4 and, in some cases, 0.5.[9] There appears to be a cross-sectional correlation between $\delta$ and $k$: those stocks that have higher $k$'s also often have higher $\delta$'s.

It is straightforward to demonstrate that the unconditional standard deviation of volatility, denoted by $s_u(\sigma)$, is given by

$$s_u(\sigma) = k/(2\delta)^{1/2}. \tag{17}$$

This formula allows one to calculate the unconditional probability of obtaining values of $\sigma$ less than zero with the arithmetic AR1 process.

[7] It should be noted that the implied volatilities used in these studies are generated from pricing models that assume nonstochastic volatility. If volatility is in reality stochastic, then the implied volatilities will be subject to measurement error, and any parameter estimates derived from them may be biased. However, as Stein (1989) argues, any such biases are likely to be extremely small. This is because variations in measurement errors for a given option are dwarfed by variations in the level of volatility. (Table 1 provides some intuition for how the measurement error on a given option changes with a change in the level of volatility.) Moreover, the parameter values used below are only intended to give a rough, "ballpark" idea of the relevant magnitudes. We do not intend to suggest that they represent the best possible statistical estimates.

[8] These estimates of the half-life of stock index volatility are broadly consistent with those seen in a number of other studies using a variety of other empirical formulations.

[9] Merville and Pieptea (1989) do not present $k$ directly, since they are not exactly estimating an AR1 process. However, it is straightforward to recombine their parameter estimates to calculate what $k$ would have been had they specified their empirical model as an AR1. In particular, $k = \sigma_D/(1 - \mathrm{NVR})^{1/2}$, where the variables are defined in their table 4 (p. 205).

For example, a typical set of parameter values for an individual stock would be $\theta = 0.30$, $\delta = 16$, $k = 0.4$. This implies that the standard deviation of volatility is 0.07, or less than one fourth of its mean. Clearly, the probability of observing a negative $\sigma$ is extremely small. Other configurations of the parameters lead to similar conclusions, as it is difficult to generate reasonable examples where the unconditional probability of a negative $\sigma$ exceeds 1 or 2 percent.

## 2.2 Computational considerations

The techniques involved in computing prices numerically based on the pricing formulas are neither complex nor particularly costly in terms of computer resources. The code used to produce the results presented here, essentially a collection of numerical integration routines, was written by Ron Henderson in the "C" programming language and was implemented by him on a Silicon Graphics Iris 4D 240 GTX workstation. At the heart of the computations is a Romberg integration routine with some slight modifications to truncate an integral over infinity to some finite region containing enough information to produce accurate results. Romberg integration was chosen over several other candidate schemes (i.e., quadrature formulas) because of its robustness and ability to adjust to a widely varying integrand.

Our program was able to generate options prices for most parameter values in less than one minute, and in many cases, in less than 15 seconds. Using the approximate distributions instead of the exact ones reduced the amount of computation by a factor that varied between about 10 and 100. It may also be possible to streamline the computation of prices based on the exact distributions. One alternative to the current approach would be to use a fast Fourier transform (FFT) in the calculations. This would effectively reduce the two integrations now involved to a single real FFT and a summation, and could bring computational "cost" to a level comparable to that required to evaluate the Black–Scholes formula numerically. We are currently pursuing this possibility.

## 2.3 Sample options prices

Options prices based on the exact stock price distributions $S(P, t)$ are presented in Table 1. The table is divided into 10 panels, labeled A–J. Each panel looks at seven strike prices and three maturities (one month, three months, and six months)—a total of 21 options. For each option, three numbers are calculated: the Black–Scholes price, a "new" price corresponding to our exact distribution, and the Black–Scholes implied volatility associated with the new price.

The panels cover a broad range of the parameter values discussed in Section 2.1. [In all cases, we set $\phi = 0$ and $r = \ln(1.1) = 9.53$

**Table 1**
**Comparison of Black-Scholes and new prices**

| | 1 Month | | | 3 Months | | | 6 Months | | |
|---|---|---|---|---|---|---|---|---|---|
| Strike | Black-Scholes | New | Implied vol. | Black-Scholes | New | Implied vol. | Black-Scholes | New | Implied vol. |
| A. | ($\sigma = 0.20$, $\theta = 0.20$, $\delta = 4.00$, $k = 0.10$) | | | | | | | | |
| 90 | 10.76 | 10.77 | 0.2023 | 12.55 | 12.57 | 0.2024 | 15.12 | 15.16 | 0.2025 |
| 95 | 6.20 | 6.20 | 0.2009 | 8.50 | 8.52 | 0.2013 | 11.34 | 11.38 | 0.2017 |
| 100 | 2.71 | 2.71 | 0.2003 | 5.23 | 5.24 | 0.2008 | 8.14 | 8.18 | 0.2013 |
| 105 | 0.83 | 0.84 | 0.2006 | 2.90 | 2.91 | 0.2008 | 5.58 | 5.62 | 0.2012 |
| 110 | 0.17 | 0.18 | 0.2015 | 1.44 | 1.46 | 0.2013 | 3.66 | 3.69 | 0.2013 |
| 115 | 0.02 | 0.03 | 0.2029 | 0.65 | 0.67 | 0.2021 | 2.29 | 2.33 | 0.2017 |
| 120 | 0.00 | 0.00 | 0.2050 | 0.26 | 0.28 | 0.2032 | 1.38 | 1.42 | 0.2022 |
| B. | ($\sigma = 0.20$, $\theta = 0.20$, $\delta = 8.00$, $k = 0.20$) | | | | | | | | |
| 90 | 10.76 | 10.77 | 0.2070 | 12.55 | 12.60 | 0.2060 | 15.12 | 15.20 | 0.2056 |
| 95 | 6.20 | 6.21 | 0.2029 | 8.50 | 8.55 | 0.2037 | 11.34 | 11.44 | 0.2044 |
| 100 | 2.71 | 2.72 | 0.2011 | 5.23 | 5.28 | 0.2025 | 8.14 | 8.24 | 0.2037 |
| 105 | 0.83 | 0.85 | 0.2020 | 2.90 | 2.95 | 0.2025 | 5.58 | 5.68 | 0.2035 |
| 110 | 0.17 | 0.19 | 0.2048 | 1.44 | 1.50 | 0.2035 | 3.66 | 3.76 | 0.2037 |
| 115 | 0.02 | 0.03 | 0.2088 | 0.65 | 0.70 | 0.2054 | 2.29 | 2.40 | 0.2043 |
| 120 | 0.00 | 0.00 | 0.2133 | 0.26 | 0.31 | 0.2077 | 1.38 | 1.48 | 0.2052 |
| C. | ($\sigma = 0.20$, $\theta = 0.20$, $\delta = 16.00$, $k = 0.30$) | | | | | | | | |
| 90 | 10.76 | 10.78 | 0.2103 | 12.55 | 12.61 | 0.2073 | 15.12 | 15.22 | 0.2066 |
| 95 | 6.20 | 6.23 | 0.2046 | 8.50 | 8.57 | 0.2051 | 11.34 | 11.46 | 0.2058 |
| 100 | 2.71 | 2.73 | 0.2021 | 5.23 | 5.31 | 0.2041 | 8.14 | 8.28 | 0.2053 |
| 105 | 0.83 | 0.86 | 0.2033 | 2.90 | 2.98 | 0.2041 | 5.58 | 5.73 | 0.2052 |
| 110 | 0.17 | 0.20 | 0.2073 | 1.44 | 1.52 | 0.2050 | 3.66 | 3.80 | 0.2053 |
| 115 | 0.02 | 0.04 | 0.2127 | 0.65 | 0.72 | 0.2067 | 2.29 | 2.43 | 0.2057 |
| 120 | 0.00 | 0.00 | 0.2187 | 0.26 | 0.32 | 0.2088 | 1.38 | 1.50 | 0.2063 |
| D. | ($\sigma = 0.25$, $\theta = 0.25$, $\delta = 4.00$, $k = 0.20$) | | | | | | | | |
| 90 | 10.88 | 10.90 | 0.2550 | 13.03 | 13.10 | 0.2558 | 15.94 | 16.05 | 0.2565 |
| 95 | 6.56 | 6.58 | 0.2521 | 9.26 | 9.32 | 0.2536 | 12.44 | 12.56 | 0.2550 |
| 100 | 3.28 | 3.29 | 0.2510 | 6.19 | 6.24 | 0.2525 | 9.45 | 9.56 | 0.2541 |
| 105 | 1.31 | 1.33 | 0.2515 | 3.88 | 3.93 | 0.2525 | 6.99 | 7.10 | 0.2538 |
| 110 | 0.41 | 0.43 | 0.2534 | 2.28 | 2.35 | 0.2535 | 5.03 | 5.15 | 0.2541 |
| 115 | 0.10 | 0.12 | 0.2563 | 1.26 | 1.33 | 0.2552 | 3.54 | 3.66 | 0.2548 |
| 120 | 0.02 | 0.03 | 0.2598 | 0.66 | 0.73 | 0.2574 | 2.42 | 2.56 | 0.2559 |
| E. | ($\sigma = 0.25$, $\theta = 0.25$, $\delta = 8.00$, $k = 0.30$) | | | | | | | | |
| 90 | 10.88 | 10.91 | 0.2589 | 13.03 | 13.13 | 0.2585 | 15.94 | 16.09 | 0.2588 |
| 95 | 6.56 | 6.59 | 0.2539 | 9.26 | 9.36 | 0.2558 | 12.44 | 12.61 | 0.2574 |
| 100 | 3.28 | 3.30 | 0.2520 | 6.19 | 6.28 | 0.2545 | 9.45 | 9.63 | 0.2566 |
| 105 | 1.31 | 1.34 | 0.2530 | 3.88 | 3.97 | 0.2545 | 6.99 | 7.17 | 0.2563 |
| 110 | 0.41 | 0.45 | 0.2563 | 2.28 | 2.38 | 0.2557 | 5.03 | 5.22 | 0.2566 |
| 115 | 0.10 | 0.13 | 0.2610 | 1.26 | 1.37 | 0.2578 | 3.54 | 3.72 | 0.2572 |
| 120 | 0.02 | 0.03 | 0.2665 | 0.66 | 0.76 | 0.2604 | 2.42 | 2.61 | 0.2582 |
| F. | ($\sigma = 0.25$, $\theta = 0.25$, $\delta = 16.00$, $k = 0.40$) | | | | | | | | |
| 90 | 10.88 | 10.92 | 0.2608 | 13.03 | 13.13 | 0.2587 | 15.94 | 16.09 | 0.2587 |
| 95 | 6.56 | 6.61 | 0.2553 | 9.26 | 9.37 | 0.2567 | 12.44 | 12.62 | 0.2579 |
| 100 | 3.28 | 3.31 | 0.2530 | 6.19 | 6.30 | 0.2557 | 9.45 | 9.65 | 0.2575 |
| 105 | 1.31 | 1.35 | 0.2541 | 3.88 | 4.00 | 0.2558 | 6.99 | 7.20 | 0.2573 |
| 110 | 0.41 | 0.46 | 0.2579 | 2.28 | 2.40 | 0.2566 | 5.03 | 5.24 | 0.2575 |
| 115 | 0.10 | 0.14 | 0.2632 | 1.26 | 1.38 | 0.2581 | 3.54 | 3.74 | 0.2578 |
| 120 | 0.02 | 0.04 | 0.2693 | 0.66 | 0.76 | 0.2601 | 2.42 | 2.61 | 0.2584 |

**Table 1**
**Continued**

| | 1 Month | | | 3 Months | | | 6 Months | | |
|---|---|---|---|---|---|---|---|---|---|
| Strike | Black–Scholes | New | Implied vol. | Black–Scholes | New | Implied vol. | Black–Scholes | New | Implied vol. |
| G. | ($\sigma = 0.35, \theta = 0.35, \delta = 4.00, k = 0.40$) | | | | | | | | |
| 90 | 11.34 | 11.39 | 0.3588 | 14.31 | 14.48 | 0.3619 | 17.93 | 18.25 | 0.3650 |
| 95 | 7.44 | 7.48 | 0.3545 | 10.95 | 11.10 | 0.3587 | 14.83 | 15.15 | 0.3628 |
| 100 | 4.42 | 4.45 | 0.3529 | 8.13 | 8.27 | 0.3572 | 12.12 | 12.43 | 0.3616 |
| 105 | 2.36 | 2.40 | 0.3537 | 5.87 | 6.01 | 0.3572 | 9.80 | 10.11 | 0.3612 |
| 110 | 1.13 | 1.18 | 0.3565 | 4.12 | 4.28 | 0.3586 | 7.83 | 8.16 | 0.3616 |
| 115 | 0.49 | 0.54 | 0.3606 | 2.82 | 3.00 | 0.3610 | 6.20 | 6.55 | 0.3626 |
| 120 | 0.19 | 0.24 | 0.3657 | 1.88 | 2.08 | 0.3642 | 4.87 | 5.23 | 0.3641 |
| H. | ($\sigma = 0.35, \theta = 0.35, \delta = 8.00, k = 0.50$) | | | | | | | | |
| 90 | 11.34 | 11.41 | 0.3612 | 14.31 | 14.50 | 0.3629 | 17.93 | 18.25 | 0.3650 |
| 95 | 7.44 | 7.50 | 0.3516 | 10.95 | 11.12 | 0.3602 | 14.83 | 15.17 | 0.3636 |
| 100 | 4.42 | 4.47 | 0.3540 | 8.13 | 8.30 | 0.3588 | 12.12 | 12.47 | 0.3629 |
| 105 | 2.36 | 2.41 | 0.3550 | 5.87 | 6.04 | 0.3588 | 9.80 | 10.15 | 0.3626 |
| 110 | 1.13 | 1.20 | 0.3584 | 4.12 | 4.31 | 0.3600 | 7.83 | 8.20 | 0.3629 |
| 115 | 0.49 | 0.56 | 0.3634 | 2.82 | 3.02 | 0.3622 | 6.20 | 6.57 | 0.3635 |
| 120 | 0.19 | 0.25 | 0.3694 | 1.88 | 2.10 | 0.3649 | 4.87 | 5.24 | 0.3645 |
| I. | ($\sigma = 0.35, \theta = 0.35, \delta = 16.00, k = 0.60$) | | | | | | | | |
| 90 | 11.34 | 11.41 | 0.3615 | 14.31 | 14.48 | 0.3616 | 17.93 | 18.20 | 0.3628 |
| 95 | 7.44 | 7.51 | 0.3567 | 10.95 | 11.12 | 0.3599 | 14.83 | 15.13 | 0.3622 |
| 100 | 4.42 | 4.48 | 0.3548 | 8.13 | 8.31 | 0.3592 | 12.12 | 12.44 | 0.3618 |
| 105 | 2.36 | 2.42 | 0.3557 | 5.87 | 6.05 | 0.3592 | 9.80 | 10.13 | 0.3617 |
| 110 | 1.13 | 1.20 | 0.3589 | 4.12 | 4.31 | 0.3599 | 7.83 | 8.17 | 0.3618 |
| 115 | 0.49 | 0.56 | 0.3637 | 2.82 | 3.01 | 0.3611 | 6.20 | 6.53 | 0.3621 |
| 120 | 0.19 | 0.25 | 0.3694 | 1.88 | 2.06 | 0.3628 | 4.87 | 5.19 | 0.3626 |
| J. | ($\sigma = 0.35, \theta = 0.25, \delta = 16.00, k = 0.40$) | | | | | | | | |
| 90 | 11.34 | 11.14 | 0.3129 | 14.31 | 13.42 | 0.2832 | 17.93 | 16.33 | 0.2715 |
| 95 | 7.44 | 7.07 | 0.3097 | 10.95 | 9.78 | 0.2817 | 14.83 | 12.93 | 0.2709 |
| 100 | 4.42 | 3.95 | 0.3085 | 8.13 | 6.79 | 0.2811 | 12.12 | 10.00 | 0.2705 |
| 105 | 2.36 | 1.92 | 0.3091 | 5.87 | 4.50 | 0.2811 | 9.80 | 7.56 | 0.2704 |
| 110 | 1.13 | 0.82 | 0.3111 | 4.12 | 2.85 | 0.2817 | 7.83 | 5.61 | 0.2705 |
| 115 | 0.49 | 0.32 | 0.3143 | 2.82 | 1.74 | 0.2828 | 6.20 | 4.08 | 0.2708 |
| 120 | 0.19 | 0.11 | 0.3182 | 1.88 | 1.02 | 0.2842 | 4.87 | 2.91 | 0.2713 |

The "new" price corresponds to the equations

$$dP = \mu P\,dt + \sigma P\,dz_1 \tag{1}$$

and

$$d\sigma = -\delta(\sigma - \theta)\,dt + k\,dz_2. \tag{2}$$

For all entries, $P = 100$, the riskless rate $r = 9.53$ percent, and the volatility risk premium $\phi = 0$. The Black–Scholes price corresponds to the nonstochastic volatility setting where $\delta = k = 0$.

percent.] Panels A–C encompass the values that appear to characterize stock index options, setting $\theta = \sigma_0 = 0.20$ and allowing $\delta$ to range from 4 to 16, while $k$ ranges from 0.10 to 0.30. The subsequent panels examine higher values of volatility and $k$ that seem appropriate for options on individual stocks. In D–F, $\theta = \sigma_0 = 0.25$ and $k$ ranges from 0.20 to 0.40. In G–I, $\theta = \sigma_0 = 0.35$ and $k$ ranges from 0.40 to 0.60. Finally, in panel J, we consider a case where initial volatility differs from its long-run mean, replicating all the parameters of panel F except setting $\sigma_0 = 0.35$, rather than its long-run mean of 0.25.

Several observations emerge from the table. First, stochastic volatility exerts an upward influence on all options prices. Whenever $\sigma_0 = \theta$, the new price exceeds the Black–Scholes price for the same $\theta$. Second, stochastic volatility is "more important" for away-from-the-money options than for at-the-money options, in the sense that the implied volatilities corresponding to the new prices exhibit a U-shape as the strike price is varied. Implied volatility is lowest at-the-money, and rises as the strike price moves in either direction.

The concept of the mixing distribution for $\sigma$ provides a useful heuristic device for understanding these effects. Given Equation (11), one can always represent our "new" price as an average of Black–Scholes prices with different $\sigma$'s, weighted by the mixing distribution $m_t(\sigma)$. Intuitively, the difference between the Black–Scholes and new prices should depend both on the mean of the distribution $m_t(\sigma)$ as well as on its dispersion.

As it turns out, the Black–Scholes formula is very close to linear in volatility for at-the-money options. This suggests that, for these options, all that should matter (loosely speaking) is the *mean* of the mixing distribution. Now, even when $\sigma_0 = \theta$, it is *not* the case that the mean of the mixing distribution equals $\theta$. For example, when volatility evolves deterministically over time, the mean of the mixing distribution is given by the square root of the average value of $\sigma^2$ over the life of the option [see Equation (A2) in Appendix A]. By Jensen's inequality, this is greater than the average value of $\sigma$. A similar (though more complex) logic also applies for the case when volatility is stochastic. This "mean of the mixing distribution effect" (roughly) explains the implied volatilities seen at-the-money.

For away-from-the-money options, there is a second effect. For these options, the Black–Scholes formula is convex in volatility. Thus, for a fixed *mean* of the mixing distribution, these options are more valuable when the mixing distribution has more dispersion. This "dispersion of the mixing distribution effect" explains the U-shape in implied volatilities mentioned above.

Table 1 shows that the overall impact on options prices can be economically significant, especially when the options are out-of-the-money and the parameter $k$ is allowed to take on large (but plausible) values. For example, in panel H, where $\sigma_0 = \theta = 0.35$, $\delta = 8$, and $k = 0.50$, our model prices a three-month option with a 120 strike at 2.10, or 11.7 percent more than its Black–Scholes price of 1.88. The new price corresponds to an implied volatility of 36.5 percent. Even larger proportional effects can be observed with cheaper options. A one-month option with a 120 strike has a new price of 0.25, which is 31.6

**Table 2**
**Comparison of exact and approximate prices**

($\sigma = 0.25$, $\theta = 0.25$, $\delta = 16.00$, $k = 0.40$)

| Strike | Black–Scholes | New | Approximation |
|---|---|---|---|
| 1 Month | | | |
| 90 | 10.88 | 10.92 | 10.91 |
| 95 | 6.56 | 6.61 | 6.61 |
| 100 | 3.28 | 3.31 | 3.34 |
| 105 | 1.31 | 1.35 | 1.37 |
| 110 | 0.41 | 0.46 | 0.45 |
| 115 | 0.10 | 0.14 | 0.12 |
| 120 | 0.02 | 0.04 | 0.03 |
| 3 Months | | | |
| 90 | 13.03 | 13.13 | 13.13 |
| 95 | 9.26 | 9.37 | 9.39 |
| 100 | 6.19 | 6.30 | 6.34 |
| 105 | 3.88 | 4.00 | 4.04 |
| 110 | 2.28 | 2.40 | 2.43 |
| 115 | 1.26 | 1.38 | 1.38 |
| 120 | 0.66 | 0.76 | 0.75 |
| 6 Months | | | |
| 90 | 15.94 | 16.09 | 16.10 |
| 95 | 12.44 | 12.62 | 12.65 |
| 100 | 9.45 | 9.65 | 9.68 |
| 105 | 6.99 | 7.20 | 7.23 |
| 110 | 5.03 | 5.24 | 5.28 |
| 115 | 3.54 | 3.74 | 3.76 |
| 120 | 2.42 | 2.61 | 2.63 |

The "new" price corresponds to the equations

$$dP = \mu P\,dt + \sigma P\,dz_1 \tag{1}$$

and

$$d\sigma = -\delta(\sigma - \theta)\,dt + k\,dz_2. \tag{2}$$

For all entries, $P = 100$, the riskless rate $r = 9.53$ percent, and the volatility risk premium $\phi = 0$. The "approximate" price corresponds to the approximation technique described in Section 1.2 of the text. The Black–Scholes price corresponds to the nonstochastic volatility setting where $\delta = k = 0$.

percent more than its Black–Scholes price of 0.19, and which corresponds to an implied volatility of 36.9 percent.

In Table 2, the approximate prices [based on the approximate distribution $\hat{S}(P, t)$] are compared to the exact new prices. Black–Scholes prices are also included as a benchmark for comparison.[10] In Table 2, the same parameter values are used as in panel F of Table 1, and the results are representative of those seen with other parameter values. As the table shows, the approximate prices are quite close to the exact prices for away-from-the-money options. For example, at a maturity of three months and a strike price of 120, the approximate

[10] By comparing the errors incurred with our approximation technique to the errors incurred with the Black–Scholes formula, one can get a rough idea of how useful the approximation technique is relative to the "default" option of not modeling stochastic volatility at all.

price is 0.75, as compared to an exact price of 0.76 and a Black–Scholes price of 0.66. In other words, the approximation error is roughly one tenth of the error incurred in using Black–Scholes.

According to this criterion, the approximation technique works somewhat less well at-the-money, although it still outperforms Black–Scholes substantially. At a maturity of three months and a strike price of 100, the approximate price is 6.34, as compared to an exact price of 6.30 and a Black–Scholes price of 6.19. In this case, the approximation error is about one third of the error incurred with Black–Scholes.[11]

## 3. Asymptotic Behavior of Stock Price Distributions

In this section, we explore the connection between the parameters of the process driving $\sigma$ and the degree to which stock price distributions have fat tails. In discussing fat tails, we focus primarily on the asymptotic shape of stock price distributions, and on the related question of what moments of the distribution exist.[12] We begin by stating the following definition.

*Definition.* Two functions $F(z)$ and $G(z)$ are *asymptotically equivalent* as $z \to \infty$ (or as $z \to 0$) if $\log F(z)/\log G(z) \to 1$ as $z \to \infty$ (or as $z \to 0$). This will be denoted as $F(z) \simeq G(z)$.

Because we are looking at asymptotic behavior, the reflecting barrier assumption inherent in the arithmetic process of Equation (2) will not color our conclusions. The existence of a reflecting barrier at $\sigma = 0$ may lead to unnatural implications about the movements of $\sigma$ when it is close to zero. However, all that is important for the asymptotics is the nature of $\sigma$'s movements when it is large (i.e., how quickly can $\sigma$ move toward infinity?).

To see this point heuristically, note that we are looking at processes for $\sigma$ of the form

$$d\sigma = a(\sigma)\ dt + b(\sigma)\ dz. \tag{18}$$

The asymptotic nature of stock price distributions will be determined

[11] The fact that there are larger errors at-the-money suggests that our approximate distribution does a better job of matching the tails of the true distribution than it does of matching the central part of the true distribution. Clearly, a more complete analysis would involve a detailed comparison (perhaps via simulation) of the approximate and true distributions. Our aim is not so much to make a strong case for the use of the particular approximation described here, especially since the exact formula is relatively easily implemented. Rather, we believe that the general method of approximation is instructive—as we explain in Section 4, it may be of practical relevance in more complicated models whose exact solutions prove elusive.

[12] Empirical studies often quantify fat tails by computing an estimate of the kurtosis, or fourth moment of the distribution. Unfortunately, such higher moments do not always converge for our theoretical stochastic volatility distributions, so we are unable to make a direct comparison in this regard.

by the limiting behavior of $a(\sigma)$ and $b(\sigma)$ as $\sigma$ gets large. Thus, the noteworthy difference between an arithmetic and a geometric Brownian motion model for $\sigma$ is that in the former $b(\sigma)$ remains constant as $\sigma$ gets larger, while in the latter it increases indefinitely with $\sigma$. The fact that $b(\sigma)$ is also constant for small $\sigma$ may be an unrealistic aspect of arithmetic Brownian motion, but it does not affect asymptotic stock price distributions—a more reasonable process that had $b(\sigma)$ shrinking for small $\sigma$ but remaining bounded for large $\sigma$ would lead to the same basic conclusions.

Once the stock price distributions are given, their asymptotic order can be recovered using a well-known technique called Laplace's method, or the theory of stationary real phase. [For a complete description, see Erdelyi (1956, pp. 36–38).] The derivations are outlined in Appendix D, where we also make the heuristic argument above more precise. Here we simply present and discuss our results.

The benchmark for comparison is the lognormal distribution $L(\sigma)$. Its asymptotic order is given by

$$L(\sigma) \simeq \exp(-(\log P)^2/2\sigma^2 t) \qquad \text{as } P \to 0 \text{ or } P \to \infty. \tag{19}$$

Thus, the relative thinness of the tails of the lognormal is reflected in the rapid decrease of the exponential as $P$ goes to zero or infinity.

In comparison, the stochastic volatility distributions studied here have asymptotic behavior that can be written as

$$S(P, t) \simeq P^{-\gamma} \qquad \text{as } P \to \infty, \tag{20}$$

$$S(P, t) \simeq P^{-1+\gamma} \qquad \text{as } P \to 0. \tag{21}$$

The exponent $\gamma$ is given by

$$\gamma = \tfrac{3}{2} + \tfrac{1}{2}(1 + 4/t\bar{t})^{1/2}, \tag{22}$$

where

$$\bar{t} = k^2 t/(v^2 + t^2\delta^2).$$

and $v = v(t\delta)$ is the smallest positive root of the equation $\cos v + (t\delta/v)\sin v = 0$. The variable $v$ always lies between $\pi/2$ and $\pi$, and $v = \pi/2$ when $\delta = 0$.

Several observations follow from (20)–(22). First, and most obviously, the asymptotic behavior of stock price distributions, which might be termed "power behavior," implies slower rates of decrease (i.e., fatter tails) than the lognormal. This power behavior is compatible with the generalized beta distribution (GB2) introduced by Bookstaber and McDonald (1987) to fit stock return data. In particular, an inspection of their formula (1a) (p. 403) shows that our parameter $\gamma$ is essentially equivalent to unity plus the product of their parameters

$a$ and $q$. They note that no moments of order equal to or higher than $aq$ will exist; analogously, it can be shown in our models that no moments of order equal to or greater than $(\gamma - 1)$ will exist for stock price distributions. For example, when $\gamma \leq 3$, the distribution will not have a well-defined variance.

Next, it is easy to verify that in our model $\gamma$ always exceeds 2, and can exceed 3, depending on the values of the parameters. Thus, for this model, stock prices sometimes have a well-defined variance.[13]

Equation (22) implies that $\gamma$ approaches infinity as $t$ or $k$ goes to zero. When $\sigma$ follows an arithmetic process, decreasing its end-of-horizon variance makes stock price tails thinner, and higher and higher moments exist.

If $\delta = 0$, there is no mean reversion in $\sigma$, and $\gamma$ approaches 2 as $t$ or $k$ goes to infinity. Thus, if the horizon is very long, so that there is a great deal of variance in the ultimate value of $\sigma$, the stock price distribution gets very fat-tailed, losing all its higher order moments up to and including the variance.

In contrast, when there is a nonzero mean reversion coefficient $\delta$, $\gamma$ approaches a number that is strictly greater than 2 as $t$ goes to infinity. The greater is $\delta$, the larger is this limiting $\gamma$. Intuitively, a nonzero $\delta$ bounds the end-of-horizon variance for $\sigma$ away from infinity, no matter how long the horizon. Consequently, the limiting case of $\gamma = 2$ is never approached, even for very large values of $t$.

The implications of the model with a positive $\delta$ appear to accord closely with the empirical findings of Bookstaber and McDonald (1987). They note that while one- and five-day stock returns have significantly fatter tails than lognormals of the same variance, returns over longer horizons (e.g., 250 days) are much better described by a lognormal distribution. This observation, taken together with our analytical results, would seem to provide indirect support for the hypothesis that volatility follows a stationary process.[14] If volatility were nonstationary, then our results would lead one to expect long-horizon returns that look substantially fatter-tailed than lognormals.

## 4. Conclusions and Extensions

We have used analytic techniques to derive both exact and approximate stock price distributions for the case where stock price dynamics are given by Equations (1) and (2). Our results have enabled us

[13] In contrast, we can show that a geometric process for $\sigma$ leads to "hyper-fat" tails—a $\gamma$ of 2 and an unbounded variance for any nonzero values of $t$ and $k$.

[14] The question of whether volatility contains a unit root has been the subject of a great deal of direct testing. Schwert (1987) provides a detailed discussion of the issues that arise in such direct testing for stationarity.

to develop closed-form options pricing formulas that incorporate important aspects of the time-series properties of volatility, as well as to sketch some links between these time-series properties and the extent to which stock price distributions have fat tails.

Although our approximation technique appears to work relatively well in pricing options (particularly those away from the money), one might question its usefulness, given that the exact formula can itself be quite easily implemented. However, it should be noted that our basic approximation methodology may be helpful in attacking more general models than the one studied here, where exact solutions prove less tractable.

Consider, for example, a constant elasticity of volatility (CEV) generalization of Equation (1), that is,

$$dP = rP\,dt + \sigma P^{j}\,dz_1, \tag{1'}$$

where $0 < j < 1$. This extension, when combined with Equation (2), captures other empirically relevant aspects of volatility, including the tendency for percentage returns to be more volatile when prices are low. We do not know whether a tractable exact solution exists for stock price distributions generated by (1′) and (2).

However, it would appear that we can apply a variant of our approximation technique. The logic is as follows. Suppose we know the stock price distribution corresponding to just (1′) with $\sigma$ fixed.[15] It can be shown that an analogue to Equation (11) holds in that our desired exact distribution for the stochastic volatility case can be represented as a mixture of fixed $\sigma$ CEV distributions. The mixing distribution is somewhat more complicated than $m_t(\sigma)$, but has a similar form. This suggests that even if we cannot solve for the exact distribution as readily as above, we may be able to use the same method of approximation. As in Equation (13), we might use a simple substitute for the mixing distribution to generate an approximate stock price distribution.

In this vein, it should be noted that our method of approximating $m_t(\sigma)$ can probably be refined, by allowing $\hat{m}_t(\sigma)$ to be less tightly parameterized and thereby fitting more of the characteristics of $m_t(\sigma)$. Such refinement may prove worthwhile for addressing the sorts of problems described above.

## Appendix A: Derivation of Formula for $S$

We now sketch the derivations of the results presented above. Our first observation is that the distribution of prices generated by the

[15] Cox and Ross (1976) provide a closed-form solution for this distribution in the case where $j = \frac{1}{2}$.

two stochastic equations (1) and (2) is a mixture of lognormal distributions. We begin by making this precise.

If we solve Equation (1), with $\sigma$ fixed and $\mu = 0$, the resulting distribution of prices is the lognormal $L(\sigma)$, given by

$$L(\sigma) = (2\pi t\sigma^2 P^2)^{-1/2} \exp\left(\frac{-(\log P + t\sigma^2/2)^2}{2t\sigma^2}\right). \tag{A1}$$

It is easy to show that if $\sigma$ is not constant, but a deterministic function $\sigma(t)$, the distribution of prices is given by $L(\alpha(t))$, where

$$\alpha(t) = \left(\frac{1}{t}\int_0^t \sigma^2(s)\,ds\right)^{1/2}. \tag{A2}$$

That is, in this case, prices are still lognormally distributed, with a variance that corresponds to the *average* $\sigma^2$ over the time interval.

In the case where $\sigma(t)$ is stochastic, and by Equation (2) is given by an AR1 process, we can write $\sigma = \sigma_\omega(t)$, where $\omega$ is the point in the probability space that labels the stochastic path. By the reasoning above, each path $\omega$ implies a price distribution $L(\alpha_\omega(t))$, where

$$\alpha_\omega(t) = \left(\frac{1}{t}\int_0^t \sigma_\omega^2(s)\,ds\right)^{1/2}. \tag{A3}$$

The desired price distribution $S$ is simply the expectation of $L(\alpha_\omega(t))$ so that

$$S = E_\omega\{L(\alpha_\omega(t))\}. \tag{A4}$$

We now focus on the random variable $\alpha_\omega(t)$. Let $m_t(\sigma)\,d\sigma$ be its distribution function so that

$$\mathrm{Prob}_\omega\{b > \alpha_\omega(t) > a\} = \int_a^b m_t(\sigma)\,d\sigma. \tag{A5}$$

This implies that, for any function $F$,

$$E_\omega(F(\alpha_\omega(t))) = \int F(\sigma)\,m_t(\sigma)\,d\sigma.$$

Therefore, (A4) implies

$$S = \int L(\sigma)\,m_t(\sigma)\,d\sigma. \tag{A6}$$

This is the claim that was stated in (11)—that is, our desired distribution is a mixture of lognormals, averaged via the mixing distribution $m_t(\sigma)$. This conclusion holds generally for the type of stochastic equations we consider in this article, and not just the particular example at hand [see, e.g., Hull and White (1987)].

It is clear that what we need to understand next is the mixing distribution $m_t(\sigma)$. The key to this is the formula for the "moment generating function,"

$$I(\lambda) = E_\omega(e^{-\lambda\alpha_\omega^2(t)}) = \int_0^\infty e^{-\lambda\sigma^2} m_t(\sigma)\, d\sigma,$$

which is given by the lemma below.

***Lemma.*** *For all $\lambda \geq 0$, the function $I(\lambda)$ is given by Equation (8), where the quantities appearing in its definition are given by Equations (3)–(7).*

The proof of the lemma will be described in Appendix B. In the special case where the parameters are $\sigma_0 = 0$, $\delta = 0$, and $\theta = 0$,

$$I(\lambda) = [\cosh(2\lambda)^{1/2} k t^{1/2})]^{-1/2}. \tag{A7}$$

This formula goes back to Cameron and Martin (1944).

With the lemma in hand, the exact formula can now be derived from (A1) and (A6) using the Fourier transform formula for $g(\xi)$ and the inversion formula for $f(x)$:

$$g(\xi) = \int_{-\infty}^{\infty} e^{ix\xi} f(x)\, dx, \tag{A8}$$

$$f(x) = (2\pi)^{-1} \int_{-\infty}^{\infty} e^{-ix\xi} g(\xi)\, d\xi. \tag{A9}$$

We define $f(x) = P \cdot S(P, t)$, with the change of variables $x = \log P$. Using (A6), this yields the following definition of $f(x)$:

$$f(x) = \int (2\pi t\sigma^2)^{-1/2} \exp\left(\frac{-(x + t\sigma^2/2)^2}{2t\sigma^2}\right) m(\sigma)\, d\sigma. \tag{A10}$$

Now we apply the Fourier transform formula (A8) to $f(x)$ to obtain

$$g(\xi) = \int m(\sigma) \left\{ \int (2\pi t\sigma^2)^{-1/2} \times \exp\left(\frac{-(x + t\sigma^2/2)^2}{2t\sigma^2}\right) e^{ix\xi}\, dx \right\} d\sigma. \tag{A11}$$

The term in braces is the Fourier transform of

$$(2\pi t\sigma^2)^{-1/2}\exp\left(\frac{-(x + t\sigma^2/2)^2}{2t\sigma^2}\right),$$

which equals $\exp(-(\xi^2 + i\xi)\sigma^2 t/2)$. [This is a standard fact about Fourier transforms, as illustrated in Wiener (1933, p. 50).] Therefore, Equation (A11) can be rewritten as

$$g(\xi) = \int m(\sigma)\exp\left(-(\xi^2 + i\xi)\frac{\sigma^2 t}{2}\right)d\sigma. \tag{A12}$$

Now recall that $I(\lambda)$ is defined as

$$I(\lambda) = \int_0^\infty e^{-\lambda\sigma^2}m(\sigma)\,d\sigma.$$

This definition means that (A12) can be reexpressed as

$$g(\xi) = I((\xi^2 + i\xi)t/2). \tag{A13}$$

We now apply the Fourier inversion formula (A9) to (A13) yielding

$$f(x) = (2\pi)^{-1}\int e^{-ix\xi}I\left((\xi^2 + i\xi)\frac{t}{2}\right)d\xi. \tag{A14}$$

We then make the change of variables $\xi = \eta - i/2$ in the above formula, also performing the indicated shift of contours in the complex plane. Since $\xi^2 + i\xi = \eta^2 + \frac{1}{4}$, this becomes

$$f(x) = (2\pi)^{-1}e^{-x/2}\int e^{-ix\eta}I\left(\left(\eta^2 + \frac{1}{4}\right)\frac{t}{2}\right)d\eta. \tag{A15}$$

When we recall our definitions $f(x) = P\cdot S_0(P, t)$ and $x = \log P$, Equation (A15) becomes Equation (9), which is the exact distribution formula for $S_0(P, t)$. This completes the derivation.

## Appendix B. Proof of the Lemma

According to the Feynman–Kac formula [see, e.g., Durrett (1984, pp. 229–234) and Freidlin (1985, pp. 117–126)], for suitable functions $c$ we have

$$E\left(\exp\left(\int_0^t c(\sigma(s))\,ds\right)\right) = u(\sigma_0, t), \tag{B1}$$

where $\sigma(t)$ is the AR1 process given by

$$d\sigma = -\delta(\sigma - \theta)\ dt + k\, dz_2, \tag{B2}$$

with $\sigma(0) = \sigma_0$, and $u$ is the solution of

$$\frac{1}{2}k^2\frac{\partial^2 u(x,t)}{\partial x^2} - \delta(x - \theta)\frac{\partial u(x,t)}{\partial x} + c(x)u(x,t) = \frac{\partial u(x,t)}{\partial t}, \tag{B3}$$

with the initial condition $u(x, 0) \equiv 1$.

In our case, we are dealing in effect with the situation that arises when $c(x) = -\lambda x^2$. With that choice of the function $c$, we have

$$u(\sigma_0, t) = E\left(\exp\left(-\lambda \int_0^t \sigma^2(s)\, ds\right)\right) = I(\lambda t).$$

Therefore,

$$u(\sigma_0, t) = I(\lambda t). \tag{B4}$$

To simplify the presentation we now relabel the parameters above so that our problem is reduced to that of solving the differential equation

$$\frac{1}{2}\frac{\partial^2 U(x,t)}{\partial x^2} + (Ax + B)\frac{\partial U(x,t)}{\partial x} + Cx^2 U(x,t) = \frac{\partial U(x,t)}{\partial t}, \tag{B5}$$

with the initial condition $U(x, 0) \equiv 1$.

It can be shown theoretically that this problem always has a solution of the form

$$U(x,t) = \exp(L_1 x^2/2 + M_1 x + N_1), \tag{B6}$$

where $L_1$, $M_1$, and $N_1$ are suitable functions of $t$. Once we know that $U(x, t)$ is of the above form, we can explicitly determine $L_1$, $M_1$, and $N_1$ by direct computation. The result is the following proposition.

***Proposition.*** *Equation (B5) has a solution (B6), where the functions $L_1$, $M_1$, and $N_1$ are given by the formulas (5), (6), and (7), with $k = 1$. In these formulas, we have used the definition $a = (A^2 - 2C)^{1/2}$ and $b = -A/a$.*

Finding a solution to (B5) with the functional form of (B6) is equivalent to solving three differential equations that determine $L_1$, $M_1$, and $N_1$. These are

$$\frac{1}{2}\cdot\frac{dL_1(t)}{dt} = C + \frac{1}{2}(L_1(t))^2 + AL_1(t), \tag{B7}$$

$$\frac{dM_1(t)}{dt} = L_1(t)M_1(t) + BL_1(t) + AM_1(t), \tag{B8}$$

$$\frac{dN_1(t)}{dt} = \frac{1}{2}(M_1(t))^2 + \frac{1}{2}L_1(t) + BM_1(t). \tag{B9}$$

The initial condition $U(x, 0) \equiv 1$ is equivalent to the initial conditions $L_1(0) = 0$, $M_1(0) = 0$, and $N_1(0) = 0$.

The solution of (B7) is given by (5) (with $k$ set equal to 1). Next, with $L_1(t)$ known, one solves (B8). The solution is given by (6) (again, with $k = 1$). Finally, with $L_1$ and $M_1$ known, one solves (B9), and its solution is given by (7), with $k = 1$. To check that (B6) is indeed a solution [with $L_1$, $M_1$, and $N_1$ given by (5)–(7)] is a straightforward but tedious task. It also has been checked using the computer program MATHEMATICA for symbolic manipulation.

Finally, to determine $I(\lambda)$ from these considerations, we first replace $t$ by $k^2t$ in the formulas for $L_1$, $M_1$, and $N_1$, giving us the functions $L$, $M$, and $N$, respectively. We also set $A = -\delta/k^2$, $B = \theta\delta/k^2$, and $C = -\lambda/k^2$ and, because of (B4), replace $\lambda$ by $\lambda/t$ and $x$ by $\sigma_0$. The result is the substitution (3) and the formula

$$I(\lambda) = \exp(L\sigma_0^2/2 + M\sigma_0 + N). \tag{B10}$$

The lemma stated in Appendix A is therefore proved.

## Appendix C. The Mixing Distribution and Approximate Mixing Distribution

There are three significant asymptotic characteristics of the mixing distribution $m_t(\sigma)$. The first is that

$$m_t(\sigma) \simeq e^{-\sigma^2/2\bar{t}}, \qquad \text{as} \quad \sigma \to \infty, \tag{C1}$$

where

$$\bar{t} = k^2t/(v^2 + t^2\delta^2) \tag{C2}$$

and $v = v(t\delta)$ is the smallest positive root of the equation $\cos v + (t\delta/v)\sin v = 0$ $(\pi/2 \leq v < \pi)$.

We shall prove here that if $m_t(\sigma) \simeq e^{-\alpha\sigma^2}$, for some fixed $\alpha$, as $\sigma \to \infty$, then indeed $\alpha = 1/2\bar{t}$, as claimed in (C1). To see this, consider $I(\lambda)$, which equals $\int_0^\infty e^{-\lambda\sigma^2}\cdot m_t(\sigma)\, d\sigma$, as noted above. This integral converges when $\lambda \geq 0$, and actually also does so for some negative values of $\lambda$, if $m_t(\sigma) \simeq e^{-\alpha\sigma^2}$, as $\sigma \to \infty$. In fact, the first negative value below which the integral $I(\lambda)$ diverges is exactly $\lambda = -\alpha$. Now if we examine the formula (8) [also (3)–(7)] giving the exact value of $I(\lambda)$, we see that this singularity occurs at exactly that value of $\lambda$ for which $\cosh(ak^2t) + b\sinh(ak^2t) = 0$. Now recall that $a = (A^2 - 2C)^{1/2}$, $b = -A/a$, with $A = -\delta/k^2$, $C = -\lambda/k^2t$. Making the indicated substitutions gives $\alpha = 1/2\bar{t}$, with $\bar{t}$ as in (C2).

The second important fact about the mixing distribution is that it decreases very rapidly as $\sigma \to 0$. More precisely,

$$m_t(\sigma) \leq c_1 e^{-c_2/\sigma^2}, \qquad \text{as} \quad \sigma \to 0, \tag{C3}$$

for two positive constants $c_1$ and $c_2$. This is a consequence of a corresponding rapid decrease of $I(\lambda)$ as $\lambda \to \infty$. This decrease is given by

$$I(\lambda) \leq e^{-c_3\lambda^{1/2}}, \qquad \lambda \to \infty, \tag{C4}$$

for some positive constant $c_3$. This in turn follows directly from an examination of the formula for the term $N$ entering in the definition of $I(\lambda)$. The formula for $N$ is a sum of four terms. The first and fourth terms contribute essentially $at/2 - at = -at/2$, for large values of $\lambda$, while the second and third terms contribute negligible quantities. Since $a = (A^2 - 2C)^{1/2}$, which is essentially $(2\lambda/k^2)^{1/2}$, when $\lambda$ is large, the conclusion (C4) is established.

Since

$$\int_0^\infty e^{-\lambda\sigma^2} m_t(\sigma)\, d\sigma \leq e^{-c_3\lambda^{1/2}},$$

it follows that, for each $s$,

$$\int_0^s m_t(\sigma)\, d\sigma \leq e^{\lambda s^2} e^{-c_3\lambda^{1/2}}. \tag{C5}$$

Now in the above, choose $\lambda$ so that $\lambda s^2 = \frac{1}{2}c_3\lambda^{1/2}$. Thus,

$$\int_0^s m_t(\sigma)\, d\sigma \leq e^{-(c_3/2)\lambda^{1/2}} = e^{-(c_3/2)^2 s^{-2}},$$

which asserts that the claimed estimate (C3) holds on the average, at least. The fact that the full estimate (C3) holds is proved by a more refined version of this argument.

The third important fact about the mixing distribution we want to point out is that

$$\int_0^\infty \sigma^2 m_t(\sigma)\, d\sigma = -I'(0), \tag{C6}$$

so that the mean of $\sigma^2$ with respect to the mixing distribution is easily determinable from the function $I(\lambda)$. Equation (C6) follows immediately from the definition

$$I(\lambda) = \int_0^\infty e^{-\lambda\sigma^2} m_t(\sigma)\, d\sigma.$$

The above observations concerning the mixing distribution $m_t(\sigma)$ suggest that we can approximate the distribution $m_t(\sigma)$ by a simpler one, $\hat{m}_t(\sigma)$, which has the form

$$\hat{m}_t(\sigma) = \rho e^{-\alpha\sigma^2} e^{-\beta/\sigma^2}, \tag{C7}$$

where $\alpha$, $\beta$, and $\rho$ are parameters that are picked in order to obtain the best fit with $m_t(\sigma)$.

Intuitively, the form for $\hat{m}_t(\sigma)$ is chosen because the factor $e^{-\alpha\sigma^2}$ matches the asymptotic behavior of $m_t(\sigma)$ at infinity given by (C1), and the factor $e^{-\beta/\sigma^2}$ matches the decay of $m_t(\sigma)$, as $\sigma \to 0$, given by (C3).

As we have said, we choose $\alpha = 1/2\bar{t}$, in accordance with (C1). Next, $\beta$ and $\rho$ are determined by the requirements that $\int_0^\infty \hat{m}_t(\sigma)\, d\sigma = 1$, and, like (C6), that $\int_0^\infty \sigma^2 \hat{m}_t(\sigma)\, d\sigma = -I'(0)$.

These requirements give

$$\beta = [-\alpha^{1/2} I(0) - 1/2\alpha^{1/2}]^2, \qquad \rho = (2/\sqrt{\pi}) \cdot e^{2\alpha^{1/2}\beta^{1/2}} \cdot \alpha^{1/2}. \tag{C8}$$

## Appendix D. Fat Tails

The asymptotic formula $S(P) \simeq P^{-\gamma}$, $P \to \infty$ [and the corresponding formula $S(P) \simeq P^{-1+\gamma}$, $P \to 0$] given in (20) and (21) can be derived by using Laplace's method for finding asymptotics of integrals with real phase functions. [For a description of this method, see Erdelyi (1956, pp. 36–38) and Hsu (1951).]

We use the identity (A6) and the asymptotic formula (C1). As is easily seen, this implies that $S(P, t) \simeq \bar{S}(P, t)$, where

$$\bar{S}(P, t) = \int_0^\infty e^{A(x, \sigma)} B(x, \sigma)\, d\sigma, \tag{D1}$$

and

$$A(x, \sigma) = -(x + t\sigma^2/2 - rt)^2/2t\sigma^2 - \sigma^2/2\bar{t}, \tag{D2}$$

$$B(x, \sigma) = (2\pi t\sigma^2 e^{2x} 2\pi\bar{t})^{-1/2}. \tag{D3}$$

with $x = \log P$.

We are interested in the asymptotics as $x \to \pm\infty$. Now it is not difficult to see that the main contribution to the integral (D1) when $x$ is large occurs for large values of $\sigma$ (with $|x|$ and $\sigma^2$ being roughly of the same order of magnitude). Thus, if we disregard terms of negligible size in $A(x, \sigma)$, we can simplify matters and replace $A(x, \sigma)$ by $\bar{A}(x, \sigma)$, where

$$\bar{A}(x,\sigma) = -(x + t\sigma^2/2)^2/2t\sigma^2 - \sigma^2/2\bar{t}. \tag{D4}$$

According to the recipe of stationary phase, the asymptotic behavior of this integral is given by

$$\frac{e^{\bar{A}(x,\sigma^*)}}{|\bar{A}''(x,\sigma^*)|^{1/2}} \cdot B(x,\sigma^*), \tag{D5}$$

where $\sigma^*$ is the critical point of $\bar{A}(x, \sigma)$ as a function of $\sigma$, that is,

$$\left.\frac{\partial \bar{A}(x,\sigma)}{\partial \sigma}\right|_{\sigma=\sigma^*} = 0.$$

The value of $\sigma^*$ is readily determined, and substituting it in (D5) leads to the formulas in the text.

**References**


Black, F., and M. Scholes, 1973, "The Pricing of Options and Corporate Liabilities," *Journal of Political Economy,* 81, 637–654.

Blattberg, R. C., and N. J. Gonedes, 1974, "A Comparison of the Stable and Student Distributions as Statistical Models for Stock Prices," *Journal of Business,* 47, 244–280.

Bollerslev, T., 1986, "Generalized Autoregressive Conditional Heteroskedasticity," *Journal of Econometrics,* 31, 307–328.

Bookstaber, R. M., and J. B. McDonald, 1987, "A General Distribution for Describing Security Price Returns," *Journal of Business,* 60, 401–424.

Cameron, R. H., and W. T. Martin, 1944, "The Wiener Measure of Hilbert Neighborhoods in the Space of Real Continuous Functions," *Journal of Mathematical Physics,* 34, 195–209.

Clark, P. K., 1973, "A Subordinate Stochastic Process Model with Finite Variance for Speculative Prices," *Econometrica,* 41, 135–155.

Cox, J. C., and S. A. Ross, 1976, "The Valuation of Options for Alternative Stochastic Processes," *Journal of Financial Economics,* 3, 145–166.

Cox, J. C., J. E. Ingersoll, and S. A. Ross, 1985, "An Intertemporal General Equilibrium Model of Asset Prices," *Econometrica,* 53, 363–384.

Durrett, R., 1984, *Brownian Motion and Martingales in Analysis,* Wadsworth, Belmont, Calif.

Engle, R. F., 1982, "Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation," *Econometrica,* 50, 987–1007.

Engle, R. F., and T. Bollerslev, 1986, "Modeling the Persistence of Conditional Variances," *Econometric Reviews,* 5, 1–50.

Engle, R. F., D. M. Lilien, and R. D. Robins, 1987, "Estimating Time Varying Risk Premia in the Term Structure: The ARCH-M Model," *Econometrica,* 55, 391–407.

Erdelyi, A., 1956, *Asymptotic Expansions,* Dover, New York.

Fama, E. F., 1963, "Mandelbrot and the Stable Paretian Hypothesis," *Journal of Business,* 36, 420–429.

Fama, E. F., 1965, "The Behavior of Stock Market Prices," *Journal of Business,* 38, 34–105.

Freidlin, M., 1985, *Functional Integration and Partial Differential Equations,* Princeton University Press, Princeton, N.J.

French, K. R., G. W. Schwert, and R. F. Stambaugh, "Expected Stock Returns and Volatility," *Journal of Financial Economics,* 19, 3–29.

Hsu, L. C., 1951, "On the Asymptotic Evaluation of a Class of Integrals Involving a Parameter," *American Journal of Mathematics,* 73, 625–634.

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

Johnson, H., and D. Shanno, 1987, "Option Pricing When the Variance is Changing," *Journal of Financial and Quantitative Analysis,* 22, 143–151.

Mandelbrot, B., 1963, "The Variation of Certain Speculative Prices," *Journal of Business,* 36, 394–419.

Merville, L. H., and D. R. Pieptea, 1989, "Stock Price Volatility, Mean-Reverting Diffusion, and Noise," *Journal of Financial Economics,* 24, 193–214.

Poterba, J. M., and L. H. Summers, 1986, "The Persistence of Volatility and Stock Market Fluctuations," *American Economic Review,* 76, 1142–1151.

Praetz, P. D., 1972, "The Distribution of Share Price Changes," *Journal of Business,* 45, 49–55.

Schwert, G. W., 1987, "Effects of Model Specification on Tests for Unit Roots in Macroeconomic Data," *Journal of Monetary Economics,* 20, 73–103.

Schwert, G. W., 1989, "Why Does Stock Market Volatility Change Over Time?" *Journal of Finance,* 44, 1115–1153.

Stein, J., 1989, "Overreactions in the Options Market," *Journal of Finance,* 44, 1011–1023.

Wiener, N., 1933, *The Fourier Integral and Certain of Its Applications,* Cambridge University Press, Cambridge, U.K.

Wiggins, J. B., 1987, "Option Values Under Stochastic Volatility: Theory and Empirical Estimates," *Journal of Financial Economics,* 19, 351–372.