NBER WORKING PAPER SERIES

NONPARAMETRIC ESTIMATION OF STATE-PRICE DENSITIES IMPLICIT IN FINANCIAL ASSET PRICES

Yacine Aït-Sahalia
Andrew W. Lo

Working Paper 5351

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
November 1995

We thank George Constantinides, John Heaton, Jiang Wang and especially Lars Hansen, as well as seminar participants at Duke University, Harvard University, MIT, the Econometric Society Winter Meetings (January 1995), the NBER Asset Pricing Conference (Fall 1995), the NNCM Conference (1994), and Washington University for helpful comments and discussions. Brent Andrews and John Stafford kindly provided us with S&P 500 index options data. Research support from the University of Chicago's Center for Research in Security Prices and MIT's Laboratory for Financial Engineering is gratefully acknowledged. A portion of this research was conducted during the first author's tenure as an IBM Corp. Faculty Research Fellow and the second author's tenure as an Alfred P. Sloan Research Fellow. This paper is part of NBER's research program in Asset Pricing. Any opinions expressed are those of the authors and not those of the National Bureau of Economic Research.

© 1995 by Yacine Aït-Sahalia and Andrew W. Lo. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

NBER Working Paper 5351
November 1995

NONPARAMETRIC ESTIMATION OF STATE-PRICE DENSITIES IMPLICIT IN FINANCIAL ASSET PRICES

ABSTRACT

Implicit in the prices of traded financial assets are Arrow-Debreu state prices or, in the continuous-state case, the state-price density (SPD). We construct an estimator for the SPD implicit in option prices and derive an asymptotic sampling theory for this estimator to gauge its accuracy. The SPD estimator provides an arbitrage-free method of pricing new, more complex, or less liquid securities while capturing those features of the data that are most relevant from an asset-pricing perspective, e.g., negative skewness and excess kurtosis for asset returns, volatility "smiles" for option prices. We perform Monte Carlo simulation experiments to show that the SPD estimator can be successfully extracted from option prices and we present an empirical application using S&P 500 index options.

Yacine Aït-Sahalia
Graduate School of Business
University of Chicago
1101 East 58th Street
Chicago, IL 60637-1561
and NBER

Andrew W. Lo
Sloan School of Management
MIT
50 Memorial Drive
Cambridge, MA 02142-1347
and NBER

1 Introduction

One of the most important theoretical advances in the economics of investment under uncertainty is the time-state preference model of Arrow (1964) and Debreu (1959) in which they introduce elementary securities each paying \$1 in one specific state of nature and nothing in any other state. Now known as "Arrow-Debreu" securities, they are the fundamental building blocks from which we have derived much of our current understanding of economic equilibrium in an uncertain environment. In a continuum of states, the prices of Arrow-Debreu securities are defined by the "state-price density" [SPD], which gives for each state x the price of a security paying \$1 if the state falls between x and $x + dx$.

The existence and characterization of an SPD can be obtained either in preference-based equilibrium models, e.g., Lucas (1978), Rubinstein (1976), or in the arbitrage-based models pioneered by Black and Scholes (1973) and Merton (1973). Both strands of the literature have adopted a different lexicon to denote closely related concepts. In the equilibrium framework, the SPD can be expressed in terms of a *stochastic discount factor* or *pricing kernel* such that asset prices are martingales under the actual distribution of aggregate consumption after multiplication by the stochastic discount factor [see, for example, Hansen and Jagannathan (1991) and Hansen and Richard (1987)].

In the no-arbitrage models, the SPD is often called the *risk-neutral density* based on the analysis of Cox and Ross (1976), who first observed that the Black-Scholes formula could be obtained by assuming that every investor were risk-neutral and consequently every asset in such a world must yield an expected return equal to the risk-free rate of interest. The SPD also uniquely characterizes the *equivalent martingale measure* under which all asset prices discounted at the risk-free rate of interest are martingales [see Harrison and Kreps (1979)]. Duffie (1992) uses the terminology *state-price deflator.*

Given the enormous informational content that Arrow-Debreu prices possess and the great simplification they provide for pricing complex state-contingent securities such as options and other derivatives, it is unfortunate that true Arrow-Debreu securities are not yet traded on any organized exchange and hence their prices are not directly observable.[1]

---

[1] This may change in the future with the advent of "supershares", first proposed by Garman (1978) and

However, Arrow-Debreu prices may be estimated or approximated from the prices of traded financial securities, as suggested by Banz and Miller (1978), Breeden and Litzenberger (1978) and Ross (1976). Three related approaches have been used in the literature. In the first, sufficiently strong assumptions are made on the underlying asset price dynamics for the SPD to be obtained in closed form. For example, if the asset prices follow geometric Brownian motion and the risk-free rate is constant, the SPD is log-normal—this is the Black-Scholes/Merton case. Most often the SPD cannot be computed in closed-form and numerically intensive methods must be used.[2] In the second line of research, the SPD is specified directly in some parametric form.[3] In the third approach, a prior parametric distribution is specified as a candidate SPD, typically the Black-Scholes log-normal density. The SPD is then estimated by minimizing its distance to the prior parametric distribution under the constraints that it correctly prices a selected set of derivative securities.[4]

In this paper, we propose an alternative to these three approaches in which the SPD is estimated *nonparametrically*, that is, with no parametric restrictions on either the underlying asset's price dynamics or the family of distributions that the SPD belongs to, and no need for choosing any prior distribution for the SPD.[5] While the parametric approaches are clearly preferable when the underlying asset's price process satisfies typical parametric assumptions, e.g., geometric Brownian motion, nonparametric methods are robust to violations of these assumptions. And since there is some empirical evidence that casts doubt on some of the

---

Hakansson (1976, 1977).

[2] See for example Goldenberg (1991) for formula in integral form in the case of diffusions other than geometric Brownian motion, Hull and White (1987), and Heston (1993) for an implicit characterization of the SPD in a stochastic volatility model, and Bates (1995) in a model of stochastic volatility and jumps.

[3] See Jarrow and Rudd (1982), Shimko (1993), Longstaff (1992, 1994), and Madan and Milne (1994).

[4] See Rubinstein (1994) and Jackwerth and Rubinstein (1995) who experiment with different distance criterions.

[5] In several other contexts the potential benefits of nonparametric methods for asset pricing applications have recently begun to be explored. Aït-Sahalia (1992) constructs nonparametric estimators of the diffusion process followed by the underlying asset return, as a basis to price interest rate derivatives nonparametrically. Aït- Sahalia (1994) tests parametric specifications of the spot interest rate process against a nonparametric alternative. The estimators all use discrete data, and require no discrete approximation even though the estimated model is in continuous-time. Hutchinson, Lo, and Poggio (1994) show how a neural network can approximate the Black-Scholes formula and other derivative pricing models. Boudoukh et al. (1995) price mortgages while Stutzer (1995) estimates the empirical distribution of stock returns.

more popular parametric specifications, at least for stock indexes,[6] a nonparametric SPD estimator may be more appropriate.

In particular, our nonparametric SPD estimator can yield valuable insights in at least four contexts. First, it provides us with an arbitrage-free method of pricing new, more complex, or less liquid securities, e.g., OTC derivatives or non-traded flexible options, given a subset of observed and liquid "fundamental" prices, e.g., basic call option prices, that we use to estimate the SPD.[7]

Second, our nonparametric estimator captures those features of the data that are most salient from an asset-pricing perspective and which ought to be incorporated into any successful parametric model. On the other hand, it also helps us understand what features are missed by tightly parametrized models. For example, in our empirical application to S&P 500 index options, the nonparametric SPD estimator naturally captures the so-called "volatility smile" [see Figure 2] because this is a prominent feature of the data. But we also document changes in the shape of the volatility smile over different maturities which parametric models so far have not incorporated. The nonparametric SPD estimator also exhibits persistent negative skewness and excess kurtosis [see Figure 7] because these too are features of the data. Indeed, a nonparametric analysis can often be advocated as a prerequisite to the construction of any parsimonious parametric model, precisely because important features of the data are unlikely to be missed by nonparametric estimators.

Third, and perhaps most importantly, the nonparametric estimator highlights the empirical features of the data in a way that is robust to the classical "joint hypothesis" problem that every parametric model in this literature suffers from. Our estimator is free of the typical joint hypotheses on asset price dynamics and risk premia that are typical of parametric arbitrage models, or on preferences in the equilibrium approach to derivative pricing. Of course, nonparametric techniques do require certain assumptions on the data-generating process itself, but these are typically weaker than those of parametric models and are less likely to be violated in practice.

---

[6] See, for example, Mandelbrot (1963), Fama (1965), and Lo and MacKinlay (1988).

[7] Of course, markets must be dynamically complete for such prices to be meaningful—see, for example, Constantinides (1982), and Duffie and Huang (1985). This assumption is almost always adopted, either explicitly or implicitly, in parametric derivative pricing models, and we shall adopt it here as well.

Fourth, if we make the additional assumption that underlying asset prices follow a diffusion process, we show how our estimator of the SPD can be used to estimate nonparametrically the instantaneous volatility function of the underlying asset return process.[8] Thus if we restrict attention to diffusions, we obtain the continuous-state analog to the implied binomial trees proposed by Rubinstein (1994) [see also Derman and Kani (1994) and Dupire (1994)].

In Section 2 we review the relation between SPDs and the pricing of derivative securities, in both an equilibrium and a no-arbitrage framework. Our nonparametric SPD estimator is introduced in Section 3. We present the results of extensive Monte Carlo simulation experiments in Section 4 in which we generate simulated price data under the Black-Scholes assumptions and show that our SPD estimator can successfully approximate the Black-Scholes SPD. In Section 5, we apply our SPD estimator to the pricing and delta-hedging of S&P 500 index options. Options with different times to expiration yield the family of SPDs over different horizons. We document several empirical features of the SPD over time, including the term structures of mean returns, volatility, skewness and kurtosis that are implied by these distributions. Moreover, unlike many parametric option pricing models, we show that the SPD-generated option pricing formula is capable of capturing persistent volatility "smiles" and other empirical features of market prices. We conclude in Section 6, and give technical results and details in the Appendix, including the asymptotic distribution that provides a measure of our estimator's accuracy.

2 SPDs and Derivative Securities

Since our nonparametric approach relies heavily on SPDs and their relation to both dynamic equilibrium and no-arbitrage pricing models, we start with a brief review of these results.

---

[8] This complements the approach in Aït-Sahalia (1992) where a nonparametric estimator of the same volatility function was obtained from the time series of the underlying asset returns. We show here how to estimate the volatility function from derivative prices. In particular, Girsanov's Theorem implies that they should be the same, which is a testable implication of the no-arbitrage pricing paradigm.

Readers familiar with this literature should proceed directly to Section 3.

2.1 SPDs and Dynamic Equilibria

We begin with a standard dynamic exchange economy [see Lucas (1978) and Rubinstein (1976)] in which there is a single consumption good and a representative agent who seeks to maximize his expected utility $\mathrm{E}_t\left[\int_t^T U_\tau\left(C_\tau\right)d\tau\right]$ at each date t, subject to the usual budget constraints. Under suitable assumptions for preferences and endowment shocks, it is well-known that the date-t equilibrium price S_t of a security with a single date-T liquidating payoff of $\psi(C_T)$ that is a function of aggregate consumption C_T is given by:

$$S_t \;=\; \mathrm{E}_t\Big[\psi(C_T)M_{t,T}\Big] \quad , \quad M_{t,T} \;\equiv\; \frac{U'_T(C_T)}{U'_t(C_t)}\;. \tag{2.1}$$

where $M_{t,T}$ is a stochastic discount factor.

Assuming that the conditional distribution of future consumption has a density representation $p_t(\cdot)$, we can rewrite (2.1) as:

$$\begin{aligned} \mathrm{E}_t\Big[\psi(C_T)M_{t,T}\Big] &= \int_0^\infty \psi(C_T)\frac{U'_T(C_T)}{U'_t(C_t)}p_t(C_T)dC_T && (2.2)\\ &= e^{-r_{t,\tau}\tau}\int_0^\infty \psi(C_T)p_t^*(C_T)dC_T && (2.3)\\ &= e^{-r_{t,\tau}\tau}\mathrm{E}_t^*[\psi(C_T)] && (2.4)\end{aligned}$$

where $\tau = T - t$ and

$$p_t^*(C_T) \;\equiv\; \frac{M_{t,T}p_t(C_T)}{\int_0^\infty M_{t,T}p_t(C_T)dC_T} \tag{2.5}$$

and $r_{t,\tau}$ is the continuously-compounded net rate of return between t and $T = t + \tau$ of a riskless cash account asset promising one unit of consumption at T. This version of the Euler equation shows that an asset's current price can be expressed as its discounted expected payoff, discounted at the riskless rate of interest. However, the expectation is taken with respect to the SPD p^*, a marginal-rate-of-substitution-weighted probability density function, not the original probability density function p of future consumption.

In this equilibrium framework, SPDs contain much information about preferences and asset price dynamics. For example, if parametric restrictions are imposed on the data-generating process of asset prices, e.g., geometric Brownian motion, the SPD estimator may be used to infer the preferences of the representative agent in an equilibrium model of asset prices [see, for example, Bick (1990) and He and Leland (1993)]. Alternatively, if specific preferences are imposed, e.g., logarithmic utility, the SPD may be used to infer the data-generating process of asset prices. Indeed, Rubinstein (1994) observes that any two of the following implies the third: (1) the representative agent's preferences; (2) asset price dynamics; and (3) the SPD.

2.2 SPDs in No-Arbitrage Models

In additional to the theoretical insights offered by the dynamic equilibrium framework of Section 2.1, the practical relevance of SPDs for derivative pricing and hedging applications has become apparent in "no-arbitrage" or "dynamically complete markets" models where a set of observed asset prices plays the role of state variables. For example, suppose that we observe a set of n_1 asset prices following Itô diffusions driven by n_2 independent Brownian motions:

$$dS_t = \mu_t dt + \sigma_t dW_t \tag{2.6}$$

with $n_1 \geq n_2$, and suppose that there exists a riskless asset with instantaneous rate of return r. Then path-independent derivative securities on such an asset with payoff function $\psi(S_T)$ are spanned by certain dynamic trading strategies, i.e., derivatives are redundant assets hence they may be priced by arbitrage.[9] In such applications the asset price dynamics are specified explicitly and conditions are imposed to ensure the existence of an SPD and dynamic completeness of markets [see Harrison and Kreps (1979), Duffie and Huang (1985) and Duffie (1992)].

---

[9] Additional assumptions are, of course, required such as frictionless markets, unlimited riskless borrowing and lending opportunities at the same instantaneous rate r, a known diffusion coefficient, etc. See Merton (1973, 1992) for further discussion.

In this example, the system of asset prices S in (2.6) supports an SPD if and only if the system of linear equations $\sigma_t \cdot \lambda_t = \mu_t$ admits at every date a solution λ_t such that $\exp\left[\int_t^T \lambda_\tau \cdot \lambda_\tau \, d\tau \big/ 2\right]$ has finite expectation, and $\exp\left[-\int_t^T \lambda_\tau dW_\tau - \int_t^T \lambda_\tau \cdot \lambda_\tau \, d\tau \big/ 2\right]$ has finite variance. In the presence of an SPD, markets are complete if and only if $\text{rank}(\sigma_t) = n_2$ almost everywhere. Then the SPD can be characterized explicitly without reference to preferences.

In our nonparametric approach, we do not need to assume that asset prices follow a diffusion and provide this example only as an illustrative case where primitive necessary and sufficient conditions can be given for (*i*) the existence of an SPD, and (*ii*) dynamic completeness of markets. Instead we assume (*i*) and (*ii*) directly. Denote by S_t the price of an underlying asset, $\delta_{t,\tau}$ its dividend rate between t and $T = t + \tau$, and let $p^*(S_t, S_T, \tau, r_{t,\tau}, \delta_{t,\tau})$ be the SPD of the asset price S_T, conditioned on the current price S_t. Consider now a European-style derivative security with a single liquidating payoff $\psi(S_T)$. In order to rule out arbitrage opportunities among the asset, the derivative and a risk-free cash account, the price of the derivative at t must be equal to:

$$e^{-r_{t,\tau}\tau} \int_0^{+\infty} \psi\left(S_T\right) p^*(S_t, S_T, \tau, r_{t,\tau}, \delta_{t,\tau}) dS_T. \tag{2.7}$$

For example, a European call option with maturity date T and strike price X has a payoff function $\psi(S_T) = \max(S_T - X, 0)$ hence its date-t price is simply:

$$H(S_t, X, \tau, r_{t,\tau}, \delta_{t,\tau}) \quad = \quad e^{-r_{t,\tau}\tau} \int_0^{+\infty} \max\left(S_T - K, 0\right) p^*(S_t, S_T, \tau, r_{t,\tau}, \delta_{t,\tau}) dS_T. \tag{2.8}$$

Even the most complex path-independent derivative security can be priced and hedged according to (2.7).

We propose to identify and estimate the SPD p^* from a collection of observed call option prices H. A direct computation shows that:

$$\frac{\partial H}{\partial X}\left(S_t, X, \tau, r_{t,\tau}, \delta_{t,\tau}\right) \quad = \quad e^{-r_{t,\tau}\tau} \left\{ - \int_K^{+\infty} p^*(S_t, K, \tau, r_{t,\tau}, \delta_{t,\tau}) dS_T \right\}. \tag{2.9}$$

Differentiate a second time to obtain:

$$\frac{\partial^2 H}{\partial X^2}(S_t, X, \tau, r_{t,\tau}, \delta_{t,\tau}) \quad = \quad e^{-r_{t,\tau}\tau} p^*(S_t, K, \tau, r_{t,\tau}, \delta_{t,\tau}) \ . \tag{2.10}$$

Therefore the second derivative of the call-pricing function H with respect to the strike price X, evaluated at $X = S_T$ gives us the SPD:

$$p^*(S_t, S_T, \tau, r_{t,\tau}, \delta_{t,\tau}) \quad = \quad e^{r_{t,\tau}\tau} \frac{\partial^2 H}{\partial X^2}(S_t, S_T, \tau, r_{t,\tau}, \delta_{t,\tau}) \, . \tag{2.11}$$

An alternative derivation of this result may be more revealing. As suggested by Breeden and Litzenberger (1978), consider a portfolio constructed by buying two call options struck at X and selling one struck at $X - \epsilon$ and one at $X + \epsilon$. Consider $1/(2\epsilon)$ shares of the portfolio—a "butterfly" spread, given the shape of its payoff function $\psi(S_T)$. Indeed this butterfly pays nothing outside the interval $[X - \epsilon, X + \epsilon]$. Letting ϵ tend to zero, the payoff function of the butterfly tends to a Dirac delta function with mass at X, i.e., in the limit the butterfly becomes an elementary Arrow-Debreu security paying \$1 if $S_T = X$ and nothing otherwise. The limit of its price as ϵ tends to zero should therefore be equal to $e^{-r_{t,\tau}\tau} p^*(S_t, S_T, \tau, r_{t,\tau}, \delta_{t,\tau})$. By definition of the butterfly, its price is:

$$\frac{1}{2\epsilon}\Big\{ 2H(S_t, X, \tau, r_{t,\tau}, \delta_{t,\tau}) - H(S_t, X - \epsilon, \tau, r_{t,\tau}, \delta_{t,\tau}) - H(S_t, X + \epsilon, \tau, r_{t,\tau}, \delta_{t,\tau}) \Big\} \ .$$

The limit of this expression is simply:

$$\frac{\partial^2 H}{\partial X^2}(S_t, X, \tau, r_{t,\tau}, \delta_{t,\tau})$$

hence the result follows.

For example, consider the Black-Scholes formula, corresponding to the case where S_t follows a geometric Brownian motion with constant volatility σ and the interest rate and dividend yield are also constant:

$$H_{\text{BS}}(S_t, X, \tau, r_{t,\tau}, \delta_{t,\tau}; \sigma) \quad = \quad S_t \Phi(d_1) \ - \ X e^{-r_{t,\tau}\tau} \Phi(d_2) \tag{2.12}$$

where:

$$d_1 \equiv \frac{\log(S_t/X) + (r_{t,\tau} - \delta_{t,\tau} + \frac{1}{2}\sigma^2)\tau}{\sigma\sqrt{\tau}} \tag{2.13}$$

$$d_2 \equiv \frac{\log(S_t/X) + (r_{t,\tau} - \delta_{t,\tau} - \frac{1}{2}\sigma^2)\tau}{\sigma\sqrt{\tau}} . \tag{2.14}$$

The corresponding SPD is the log-normal distribution with mean $((r_{t,\tau} - \delta_{t,\tau}) - \frac{1}{2}\sigma^2)(T-t)$ and variance $\sigma^2(T-t)$:

$$p^*_{\mathrm{BS}}(S_T, S_t, \tau, r_{t,\tau}, \delta_{t,\tau}; \sigma) = \frac{1}{S_T\sqrt{2\pi\sigma^2\tau}}\, e^{-\frac{[\log(S_T/S_t)-(r_{t,\tau}-\delta_{t,\tau}-\sigma^2/2)\tau]^2}{2\sigma^2\tau}} \tag{2.15}$$

3 Nonparametric Estimation of SPDs

We propose to estimate the SPD using (2.11). We therefore require a call option pricing formula $H(\cdot)$ that can be differentiated at least twice. Instead of relying on a particular parametric formula based on assumptions that may not be satisfied, we use observed call option prices to construct a nonparametric estimator of $H(\cdot)$.

3.1 Nonparametric Estimation of the Option Pricing Function

The specification of the typical *parametric* option pricing model $H(S_t, X, \tau, r_{t,\tau}, \delta_{t,\tau}; \theta)$ contains a vector of unknown parameter values θ which must be estimated [see, for example, Lo (1986, 1988)]. If, for example, $H(\cdot)$ is the Black-Scholes formula (2.12), θ contains only one element, the unknown volatility parameter σ. In Merton's (1976) jump-diffusion formula, θ contains four parameters.[10] To implement such parametric models, θ must be estimated either directly from observations of the underlying asset prices or indirectly from option prices themselves. For example, using option prices H_i and option characteristics $Z_i \equiv [\ S_{t_i} \ \ X_i \ \ \tau_i \ \ r_{t_i,\tau_i} \ \ \delta_{t_i,\tau_i}\ ]'$ we would typically estimate the parameter vector θ, which takes values in the finite-dimensional parameter space Θ, by solving the nonlinear least squares problem:

---

[10] Specifically, $\theta = [\ \sigma \ \ \mu \ \ \gamma \ \ \lambda\ ]'$ where σ is the coefficient of the diffusion term, λ is the Poisson jump intensity, and μ and γ are the mean and standard deviation, respectively, of the jump magnitude.

$$\min_{\theta \in \Theta} \sum_{i=1}^{n} [H_i - H(S_{t_i}, X_i, \tau_i, r_{t_i,\tau_i}, \delta_{t_i,\tau_i}; \theta)]^2. \tag{3.1}$$

However, since we do not wish to impose parametric assumptions on the data, we are faced with the more ambitious task of estimating the entire *function* H given price data for the options and their characteristics. The problem to be solved is analogous to the nonlinear least squares approach, with the major difference that the minimization is over all *functions* $H(\cdot)$ in a space of functions Γ, e.g., the space of twice continuously differentiable functions:

$$\min_{H(\cdot) \in \Gamma} \sum_{i=1}^{n} [H_i - H(S_{t_i}, X_i, \tau_i, r_{t_i,\tau_i}, \delta_{t_i,\tau_i})]^2 \tag{3.2}$$

It is well known that the solution is given by the conditional expectation of H given Z. To estimate this conditional expectation, we appeal to a statistical technique called "nonparametric kernel regression". Nonparametric kernel regression produces an estimator of the conditional expectation of H, conditioned on Z, without assuming that the function $H(\cdot)$ can be summarized by some parameter vector θ:

$$\hat{H}(S_t, X, \tau, r_{t,\tau}, \delta_{t,\tau}) \quad = \quad \hat{\mathrm{E}}[H | S_t, X, \tau, r_{t,\tau}, \delta_{t,\tau}] \tag{3.3}$$

Kernel regression requires no assumptions other than smoothness of the function to be estimated and regularity of the data used to estimate it, and is robust to the potential misspecification of any given parametric call pricing formula. On the other hand, it tends to be data-intensive. Financial applications are a natural outlet for kernel regression since typical parametric assumptions, e.g., normality or geometric Brownian motion, have been rejected by the data, and at the same time large sample sizes of high quality data are not uncommon.

A second motivation for kernel regression, one which is particularly insightful, is the local averaging or "smoothing" interpretation. Suppose that we wish to estimate the relation between two variables Z_i and H_i which satisfy the following nonlinear relation:

$$H_i \quad = \quad H(Z_i) \; + \; \epsilon_i \qquad , \qquad i = 1, \ldots, n$$

where $H(\cdot)$ is an unknown but fixed nonlinear function and $\{\epsilon_i\}$ is white noise. Now consider estimating $H(\cdot)$ at a specific point $Z_{i_0} = z_0$ and suppose for this one particular observation of Z_i, we are able to obtain *repeated* observations of the variable H_{i_0}, say $H_{i_0}^{(1)}, \ldots, H_{i_0}^{(q)}$. In this case, a natural estimator of the function $H(\cdot)$ at the point z_0 is simply:

$$\hat{H}(z_0) \quad = \quad \frac{1}{q} \sum_{j=1}^{q} H_{i_0}^{(j)} \quad = \quad H(z_0) \; + \; \frac{1}{q} \sum_{j=1}^{q} \epsilon_{i_0}^{(j)}$$

and by the law of large numbers, $\frac{1}{q} \sum_{j=1}^{q} \epsilon_{i_0}^{(j)}$ becomes negligible for large q.

Of course, if $\{H_i\}$ is a time series, we do not have the luxury of repeated observations for a given $Z_{i_0} = z_0$. However, if we assume that the function $H(\cdot)$ is smooth, then for time series observations Z_i near the value z_0, the corresponding values of H_i should be close to $H(z_0)$. In other words, if $H(\cdot)$ is smooth, then in a small neighborhood around z_0, $H(z_0)$ will be nearly constant and may be estimated by taking an average of the H_is that correspond to those Z_is near z_0. The closer the Z_is are to the value z_0, the closer an average of the corresponding H_is will be to $H(z_0)$. This argues for a *weighted* average of the H_is, where the weights decline as the Z_is get farther away from the point z_0. Such a weighted average must be computed for *each* value of z in the domain of $H(\cdot)$ to estimate the entire function, hence computational considerations become important.

This weighted average procedure of estimating $H(z)$ is the essence of smoothing. To implement such a procedure, we must define what we mean by "near" and "far". If we choose too large a neighborhood around z to compute the average, the weighted average will be too smooth and will not exhibit the genuine nonlinearities of $H(\cdot)$. If we choose too small a neighborhood around z, the weighted average will be too variable, reflecting noise as well as the variations in $H(\cdot)$. Therefore, the weights must be chosen carefully to balance these two considerations. The choice of weighting function—typically given by a probability density function [since such functions integrate to one], though the particular density function plays no probabilistic role here—determine the degree of local averaging.[11]

To specify a particular kernel regression model, we start with the natural assumption that the option pricing formula H we seek to estimate is a function of a vector of option characteristics or "explanatory" variables, $Z \equiv [\ S_t \quad X \quad \tau \quad r_{t,\tau} \quad \delta_{t,\tau}\ ]'$ so that each option price H_i, $i = 1, \ldots, n$, contained in our dataset is paired with the vector $Z_i \equiv [\ S_{t_i} \quad X_i \quad \tau_i \quad r_{t_i,\tau_i} \quad \delta_{t_i,\tau_i}\ ]'$. Since we have $d = 5$ explanatory variables, we select a five-dimensional weighting or "kernel" function $\mathrm{K}(Z)$ which integrates to one.

The density $\mathrm{K}(Z - Z_i)$, as a function of Z, has a certain spread around the data point Z_i. We can change the spread of the kernel K around Z using a "bandwidth" h, to form the new density function $(1/h)\mathrm{K}((Z - Z_i)/h)$. The closer h is to zero, the more peaked is this new density function around Z_i, and hence more weight is given to realizations of the random variable Z_i that are close to Z. An estimator (3.3) of the conditional expectation of H conditioned on Z is then given by the following expression, called the "Nadaraya-Watson" kernel estimator, where h becomes smaller as the sample size n grows:

$$\hat{H}(Z) \quad = \quad \hat{\mathrm{E}}[H|Z] \quad = \quad \frac{\sum_{i=1}^{n} \mathrm{K}((Z - Z_i)/h) H_i}{\sum_{i=1}^{n} \mathrm{K}((Z - Z_i)/h)} \; . \tag{3.4}$$

Intuitively, the estimate of the conditional expectation at a point Z, i.e., the price of an option with characteristics Z, is given by a weighted average of the observed prices H_is with more weight given to the options whose characteristics Z_is are closer to the characteristics Z of the option to be priced.

3.2 Practical Considerations: Dimension Reduction

We show in the Appendix that obtaining accurate estimates of the regression function is more difficult when the number of regressors is large [here $d = 5$], and when high-order derivatives need to be computed: recall that to obtain the SPD we differentiate the call pricing function twice.[12]

---

[11] See Härdle (1990) and Wand and Jones (1995) for a more detailed discussion of nonparametric regression.

[12] As the sample size increases, the estimator (3.4, as well as its derivatives with respect to Z, converge to the true function $H(\cdot)$ and its derivatives at every point. Therefore, when the true function satisfies certain shape restrictions, e.g., monotonicity and convexity with respect to certain variables, the estimator $\hat{H}(\cdot)$

To be fully nonparametric, all $d = 5$ regressors in $Z = [\ S_t\ \ X\ \ \tau\ \ r_{t,\tau}\ \ \delta_{t,\tau}\ ]'$ must be included in the kernel regression (3.4) of call option prices H on Z. To reduce the number of regressors, we examine the following possibilities.

First, we could assume that the option pricing formula [and hence the SPD] is not a function of the asset price S_t, the riskfree rate $r_{t,\tau}$ and the dividend yield $\delta_{t,\tau}$ separately, but only depends on these three variables through the futures price $F_{t,\tau} = S_t e^{(r_{t,\tau}-\delta_{t,\tau})\tau}$ and the riskfree rate. By no-arbitrage, the mean of the SPD depends only on—and, in fact, is equal to—$F_{t,\tau}$ and the assumption here would be that the entire distribution has this property. It is satisfied by the Black-Scholes SPD (2.15). Under this assumption, the number of regressors is reduced from $d = 5$ to $d = 4$.

In the second alternative, we could assume that the option pricing function is homogeneous of degree one in S_t and X, as in the Black-Scholes formula. This assumption would also reduce the number of regressors from $d = 5$ to $d = 4$. Combining this assumption with the previous one, the dimension of the problem would be further reduced to $d = 3$. It can be shown however that the call pricing function is homogeneous of degree one in the asset price and the strike price when the distribution of the returns is independent of the level of the asset price.[13] An example of a pricing formula satisfying the homogeneity property would be the one generated by a stochastic volatility model where the drift and diffusion functions of the stochastic volatility process depend on the volatility itself but not on the asset price. While this assumption may not be too restrictive in practice, this is nevertheless the type of assumption on the asset price dynamics that we wish to avoid in the first place by using a nonparametric estimator.

Our third proposed approach to dimension reduction is *semiparametric*. Suppose that

---

will also have these properties. However, since the asymptotic convergence of the higher-order derivatives may in practice be slow, it can be useful in small samples to modify the estimator to force it to satisfy these restrictions. One simple way to enforce monotonicity is to run an isotonic regression after the kernel regression; this guarantees that the resulting estimator $\hat{H}(\cdot)$ is monotonic [for a description of isotonic regression, see Barlow et al. (1972), and for the properties of the kernel estimator after isotonization see Mammen (1991)]. To enforce convexity in small samples, we suggest extending this procedure one step further: differentiate the isotonic kernel estimator to obtain $\hat{H}'(\cdot)$, and then run an isotonic regression on $\hat{H}'(\cdot)$. By doing so, we are guaranteed that $\hat{H}'(\cdot)$ also is monotonic, and hence $\hat{H}(\cdot)$ is convex.

[13] See Theorem 9 in Merton (1973). Merton (1973) also provides a counterexample showing that the homogeneity property can fail if the distribution of the returns is not independent of S_t.

the call pricing function is given by the parametric Black-Scholes formula (2.12) except that the volatility parameter for that option is a nonparametric function $\sigma(S_t, X, \tau)$:

$$H(S_t, X, \tau, r_{t,\tau}, \delta_{t,\tau}) \quad = \quad H_{\mathrm{BS}}(S_t, X, \tau, r_{t,\tau}, \delta_{t,\tau}; \sigma(S_t, X, \tau)) \tag{3.5}$$

In this semiparametric model, we would only need to compute the three-dimensional kernel regression of implied volatilities on asset price, strike and time-to-expiration to estimate $\mathrm{E}[\sigma|S_t, X, \tau]$. Thus $d = 3$ and the rest of the call pricing function $H(S_t, X, \tau, r_{t,\tau}, \delta_{t,\tau})$ is parametric, thereby considerably reducing the sample size n required to achieve the same degree of accuracy as the full nonparametric estimator.

3.3 The Nonparametric SPD Estimator and Its Properties

Given the observed data $\{H_i, S_{t_i}, X_i, \tau_i, r_{t_i,\tau_i}, \delta_{t_i,\tau_i}\}$, we construct the fully nonparametric call pricing function as:

$$\hat{H}(S_t, X, \tau, r_{t,\tau}, \delta_{t,\tau}) \quad = \quad \hat{\mathrm{E}}[H|S_t, X, \tau, r_{t,\tau}, \delta_{t,\tau}] \tag{3.6}$$

using in (3.4) a multivariate kernel K, formed as a product of $d = 5$ univariate kernels:

$$\hat{H}(S_t, X, \tau, r_{t,\tau}, \delta_{t,\tau}) \quad = \quad \frac{\sum_{i=1}^{n} \mathrm{k}_S(\frac{S_t - S_{t_i}}{h_S}) \mathrm{k}_X(\frac{X - X_i}{h_X}) \mathrm{k}_\tau(\frac{\tau - \tau_i}{h_\tau}) \mathrm{k}_r(\frac{r_{t,\tau} - r_{t_i,\tau_i}}{h_r}) \mathrm{k}_\delta(\frac{\delta_{t,\tau} - \delta_{t_i,\tau_i}}{h_\delta}) H_i}{\sum_{i=1}^{n} \mathrm{k}_S(\frac{S_t - S_{t_i}}{h_S}) \mathrm{k}_X(\frac{X - X_i}{h_X}) \mathrm{k}_\tau(\frac{\tau - \tau_i}{h_\tau}) \mathrm{k}_r(\frac{r_{t,\tau} - r_{t_i,\tau_i}}{h_r}) \mathrm{k}_\delta(\frac{\delta_{t,\tau} - \delta_{t_i,\tau_i}}{h_\delta})} . \tag{3.7}$$

We discuss the selection of the kernel functions and bandwidths in the Appendix. In the reduced-dimension cases, we substitute the appropriate list of regressors for $(S_t, X, \tau, r_{t,\tau}, \delta_{t,\tau})$ in (3.7). In the semiparametric model, we form the three-dimensional kernel estimator of $\mathrm{E}[\sigma|S_t, X, \tau]$ as:

$$\hat{\sigma}(S_t, X, \tau) \quad = \quad \frac{\sum_{i=1}^{n} \mathrm{k}_S(\frac{S_t - S_{t_i}}{h_S}) \mathrm{k}_X(\frac{X - X_i}{h_X}) \mathrm{k}_\tau(\frac{\tau - \tau_i}{h_\tau}) \sigma_i}{\sum_{i=1}^{n} \mathrm{k}_S(\frac{S_t - S_{t_i}}{h_S}) \mathrm{k}_X(\frac{X - X_i}{h_X}) \mathrm{k}_\tau(\frac{\tau - \tau_i}{h_\tau})} \tag{3.8}$$

where σ_i is the volatility implied by the price H_i. We then estimate the call pricing function as:

$$\hat{H}(S_t, X, \tau, r_{t,\tau}, \delta_{t,\tau}) \;=\; H_{\mathrm{BS}}(S_t, X, \tau, r_{t,\tau}, \delta_{t,\tau}; \hat{\sigma}(S_t, X, \tau)) \tag{3.9}$$

In either case, the option's delta and the SPD estimators follow by taking the appropriate partial derivatives of $\hat{H}$:

$$\hat{\Delta}(S_t, X, \tau, r_{t,\tau}, \delta_{t,\tau}) \;=\; \frac{\partial \hat{H}(S_t, X, \tau, r_{t,\tau}, \delta_{t,\tau})}{\partial S_t} \tag{3.10}$$

$$\hat{p}^*(S_t, S_T, \tau, r_{t,\tau}, \delta_{t,\tau}) \;=\; e^{r_{t,\tau}\tau}\left[\frac{\partial^2 \hat{H}(S_t, X, \tau, r_{t,\tau}, \delta_{t,\tau})}{\partial X^2}\right]_{X=S_T} . \tag{3.11}$$

3.4 Inferring Risk-Neutral Dynamics

So far, we have not assumed that the underlying asset price process S_t belonged to the class of continuous-time Itô diffusions. If we now make this additional assumption, we can derive the dynamics of the underlying asset price S_t that are compatible with our estimated SPD. In other words, we construct the risk-neutral diffusion:

$$dS_t \;=\; (r_t - \delta_t)S_t dt \;+\; \sigma(S_t)S_t dW_t \tag{3.12}$$

that would have generated the SPD that we estimated. Here r_t is the instantaneous spot interest rate [the limit of $r_{t,\tau}$ as τ approaches zero], δ_t the instantaneous dividend yield [defined similarly] and $\sigma(\cdot)$ is positive and twice-continuously-differentiable on $[0, \infty)$. Zero is an entrance barrier for the price. We assume that the interest rate and dividend yield processes, and the function $\sigma^2(\cdot)$, are such that the stochastic differential equation (3.12) admits a unique strong solution.

Constructing the risk-neutral diffusion is useful for many reasons. First, it is difficult to price American-style derivatives using the Feynman-Kac characterization (2.7) of their prices. We would need to compute the maximum value under every possible exercise policy, taking into account the potential payoffs to exercising at any date. It is often easier to price American derivatives by one of two methods.

In the first method, one constructs a binomial tree which approximates the continuous risk-neutral dynamics (3.12). Cox, Ross and Rubinstein (1979) pioneered this method for the Black-Scholes case. Nelson and Ramaswamy (1990) give a general way of constructing trees that will approximate a given diffusion. Derman and Kani (1994) build a tree which prices correctly a set of options. Rubinstein (1994) proposes a binomial tree which is not constructed form the diffusion but instead from the SPD. Our estimator of the SPD can be used directly as the SPD input of Rubinstein's tree.

The second method involves the numerical solution of a partial differential equation with free boundary to account for the possibility of early exercise. In this case as well we need the risk-neutral diffusion, since the function $\sigma^2(\cdot)$ appears as the parabolic coefficient of the partial differential equation:

$$\frac{\partial H}{\partial \tau} = \frac{\partial H}{\partial S} S(r-\delta) + \frac{1}{2}\frac{\partial^2 H}{\partial S^2} S^2 \sigma^2(S) - rH \tag{3.13}$$

Either one of these two methods requires that we estimate the risk-neutral dynamics of the asset, that is the volatility function $\sigma(\cdot)$. The function $\sigma^2(\cdot)$ can immediately be inferred from the call pricing function [and its derivatives]:

$$\sigma^2(S) = \frac{\frac{\partial H}{\partial \tau} - \frac{\partial H}{\partial S} S(r-\delta) + rH}{\frac{1}{2}\frac{\partial^2 H}{\partial S^2} S^2} . \tag{3.14}$$

We have seen earlier how to estimate the second partial derivative of the call pricing function with respect to the strike price. This formula uses other partial derivatives. They can be estimated similarly by differentiating our nonparametric estimator of the call pricing function.

4 Monte Carlo Analysis

To examine the practical performance of the nonparametric SPD estimator, we perform several Monte Carlo simulation experiments under the assumption that call option prices are truly determined by the Black-Scholes formula. Our nonparametric approach should be able to approximate Black-Scholes prices, from which the SPD may be extracted according to (2.11). The nonparametric pricing formula and SPD may then be compared to the Black-Scholes formula and theoretical SPD, respectively, to gauge the accuracy of the nonparametric approach.

Naturally, the advantage of our nonparametric approach lies in its robustness. If the options were priced by another formula, the nonparametric approach should be able to approximate it as well since, by definition, it does not rely on any parametric specification for the underlying asset's price process. Therefore, similar Monte Carlo simulation experiments can be performed for alternative option pricing models. However, we choose to perform the simulation experiments under the Black-Scholes assumptions since this is the leading case from which most applications and extensions are derived.

4.1 Calibrating the Simulations

Since our empirical application involves S&P 500 index options, we perform Monte Carlo simulation experiments to match the basic features of our dataset (see Section 5 and Table 1 for further details). We start by simulating one year, i.e., 252 days, of daily index prices generated by a geometric Brownian motion with constant drift and diffusion parameters that match the moments of the data. Specifically, the values of the initial index level, the interest rate, and the index return mean and standard deviation are fixed at 455, 3%, 7.95%, and 10.28%, respectively.[14]

At the start of this one-year sample of simulated daily prices, we create call options with strike prices and times-to-maturity that follow the Chicago Board Options Exchange [CBOE] conventions for introducing options to the market. As the index price changes from one simulated day to the next, existing options may expire and new options may be introduced, again according to CBOE conventions, with strike prices that bracket the index in 5-point increments. Therefore, on any given simulated day, the number of options is an endogenous function of the prior sample path of the index price—there are approximately

---

[14] Throughout this study we follow the common convention of reporting returns and their means and standard deviations at an annual frequency. For purposes of calculating option prices, annual parameter values are converted to a daily frequency by dividing by 365 for interest rates and returns, and by $\sqrt{252}$ for standard deviations.

80 different call options on any one simulation day, including both existing and new ones.

We then price these options by the Black-Scholes formula using the actual index volatility, and add a small white noise term to those prices. By doing so, we seek to reflect the existence in real data of a bid-ask spread and other possible sources of error in the recorded prices.[15] Since by construction the option prices in the simulation satisfy the Black-Scholes formula on average, when we apply our nonparametric pricing function to the simulated data, we should be able to "recover" the Black-Scholes formula. By this, we mean that $\hat{H}(S_t, X, \tau, r_{t,\tau}, \delta_{t,\tau})$ in (3.3) should approximate the Black-Scholes formula (2.12) numerically, not necessarily algebraically—in practice, the functional form of $\hat{H}$ may be quite different from the Black-Scholes formula (2.12), but both expressions will produce similar prices over the range of input values in the data. The two objectives of our simulation experiments is to determine how close $\hat{H}$ is to (2.12), and how close the corresponding nonparametric SPD is to the theoretical SPD (2.15).

Specifically, we take the Black-Scholes prices in the simulated dataset and the option characteristics as the inputs $\{H_i, S_{t_i}, X_i, \tau_i, r_{t_i,\tau_i}\}$ of our procedure, and then compute the smooth nonparametric call pricing function of Section 3.3. We then construct the nonparametric option delta estimator $\hat{\Delta}$ and SPD estimator $\hat{p}^*$ according to (3.10) and (3.11), respectively.

For each simulated sample path of index and options prices, we calculate several performance measures described below, and this entire procedure is repeated 5,000 times. This yields the sampling distribution for each of the performance measures and provides an indication of the accuracy of our nonparametric approach.

4.2 Accuracy of Prices, Deltas, and SPDs

To assess the performance of nonparametric option pricing formula and its corresponding delta and SPD, we first consider the percentage differences between the nonparametric option pricing formula, delta, and SPD and their theoretical Black-Scholes counterparts, respectively. In Figure 1, the theoretical values for prices, deltas, and SPDs are plotted on the left, and the average differences between the theoretical values and the nonparametric ones, averaged over the 5,000 replications, are plotted on the right. The figures show that the nonparametric quantities are within one percent of their theoretical counterparts—the estimators are virtually free of any bias. The dispersion of the estimates across the simulation runs yields the sampling distribution of the estimator. We confirmed the accuracy of the asymptotic distribution derived in the Appendix for the sample size relevant for our empirical study in Section 5.

---

[15] The white noise term is Gaussian with standard deviation equal to either one or two price ticks, depending upon whether the option characteristics made it a high- or low-volume option.

4.3 Delta-Hedging Tracking Errors

The second measure of performance is the delta-hedging "tracking error" measure proposed by Hutchinson et al. (1994), which consists of the end-of-period dollar-values of various replicating portfolios, portfolios designed to delta-hedge an option position using the nonparametric pricing function to calculate the hedge ratios or deltas.

In particular, the measure is computed by constructing the following portfolio: sell one option at date 0 and undertake the usual dynamic trading strategy in stocks and bonds to hedge this option during its life. If the option pricing model on which the delta-hedging is based is correct, and if hedging is costless and continuous, then at expiration the combined value of the stock and bond positions should exactly offset the value of the option. The difference between the terminal value of the option and the combined terminal values of the stock and bond positions then serves as a measure of the accuracy of the option pricing model. Of course, since it is impossible to hedge continuously in practice, there will always be some tracking error due to discreteness [see Boyle and Emanuel (1980) and Leland (1985) for example], therefore the SPD tracking error must be compared with the tracking error obtained from discretely-hedging with the exact Black-Scholes formula.[16]

---

[16] Formally, let V_t denote the dollar value of the replicating portfolio at date t and $V_t = V_{s,t} + V_{b,t} + V_{c,t}$ where $V_{s,t}$ is the dollar value of the underlying asset, $V_{b,t}$ is the dollar value of bonds, and $V_{c,t}$ is the dollar value of call options held in the portfolio at date t. The initial composition of this portfolio at date 0 is: $V_{s,0} = S_0\hat{\Delta}_0$, $\hat{\Delta}_0 = \frac{\partial \hat{H}_{\mathrm{BS},t=0}}{\partial S}$, $V_{c,0} = -H_{\mathrm{BS},t=0}$, $V_{b,0} = -(V_{s,0} + V_{c,0})$ where H_{BS} is the Black-Scholes call option pricing formula and $\hat{H}$ is its nonparametric estimator. Since the purchase of the underlying asset is wholly financed by the combination of riskless borrowing and proceeds from the sale of the call option, the initial value of the replicating portfolio is identically zero, that is $V_0 = 0$. Prior to expiration, and at discrete and regular intervals of length ϵ, positions in the underlying asset and bonds in the replicating portfolio will be

We use the same simulated sample paths as in the Section 4.2. Each simulated arbitrage portfolio begins with a single 252-day call option, and stock and bond holdings that yield a zero net investment for the portfolio. Portfolios are rebalanced every 5 days to delta-hedge the option [according to either the nonparametric delta or the Black-Scholes delta], and the tracking errors V_T are recorded for each of the 5,000 sample paths. Three independent sets of such simulations are performed, for strike prices of 415, 455, and 495, respectively.

Table 1 reports four performance measures based on the tracking errors V_T: the present value of the average tracking error, $e^{-rT}\mathrm{E}[V_T]$, the present value of the average absolute tracking error, $e^{-rT}\mathrm{E}[|V_T|]$, and the same two measures expressed as percentages of the initial Black- Scholes call price $H_{\mathrm{BS},0}$. The largest tracking errors occur for the out-of-the-money call, yielding an average absolute tracking error of 22.4% for the nonparametric option pricing formula. However, in this case the Black-Scholes formula yields a slightly higher average absolute tracking error, 22.7%! The fact that the Black-Scholes formula can be less accurate than the nonparametric formula may seem surprising, but recall that we are delta-hedging only once every 5 days. Despite the fact that the Black-Scholes assumptions are satisfied, there is no guarantee that delta-hedging at discrete intervals according to the Black- Scholes formula will yield smaller tracking errors than other methods.

Overall, the tracking errors of the nonparametric option pricing formulae are comparable to those of the Black-Scholes model. Of course, these results are sensitive to the parameters of the simulations and may not generalize to other parameter values. Nevertheless, these findings suggest that the nonparametric SPD estimator is reasonably accurate for the sample sizes considered in this study.

5 Extracting SPDs From S&P 500 Options

To assess the empirical relevance of our nonparametric option pricing formula and the corresponding SPD estimator, we present an application to the pricing and hedging of S&P 500

---

rebalanced so as to satisfy the following relations: $V_{s,t} = S_t\hat{\Delta}_t$, $\hat{\Delta}_t = \frac{\partial \hat{H}_t}{\partial S}$, $V_{b,t} = e^{r\epsilon}V_{b,t-\epsilon} - S_t(\hat{\Delta}_t - \hat{\Delta}_{t-\epsilon})$ where $t = q\epsilon \leq T$ for some integer q. The tracking error of the replicating portfolio is then defined to be the value of the replicating portfolio V_T at expiration date T.

index options using data obtained from the CBOE for the sample period from January 4, 1993 to December 31, 1993.

5.1 The Data

Table 2 describes the main features of our dataset. This sample contains a total of 16,923 pairs of call- and put-option prices—we take averages of bid- and ask-prices as our raw data. Observations with time-to-maturity less than one day, implied volatility greater than 70%, and price less than 1/8 are dropped, which yields a final sample of 14,431 observations and this is the starting point for our empirical analysis.

During 1993, the mean and standard deviation of continuously-compounded daily returns of the S&P 500 index was 7.95% and 10.28%, respectively [see footnote 14]. Short-term interest rates during the period exhibited little variation: they ranged from 2.85% to 3.21%. The options in our sample varied considerably in price and terms—for example, the time-to-maturity varied from 1 day to 350 days, with a median time-to-maturity of 66 days. Given the volatility and movement in the index during this period and CBOE rules for introducing new options to the market, our sample contains a fairly broad cross-section of options.

S&P 500 Index Options [symbol: SPX] are among the most actively financial derivatives in the world. Average total daily volume during the sample period was 65,476 contracts. The minimum tick for series trading below 3 is 1/16 and for all other series 1/8. Strike price intervals are 5 points, and 25 for far months. The expiration months are the three near-term months followed by three additional months from the March quarterly cycle [March, June, September, December]. The options are European, and the underlying asset is an index, the most likely case for which a lognormal assumption [with continuous dividend stream] can be justified. In other words, this market is as close as one can get to satisfying the assumptions of the Black-Scholes model. This is, therefore, a particularly promising application to test our approach: how different is our estimated SPD from the Black-Scholes SPD?

Even though the options are European and do not have a wildcard feature, the raw data present three challenges that must be addressed. First, because in-the-money options are very infrequently traded relative to at- and out-of-the-money options, in-the-money option prices are notoriously unreliable. For example, the average daily volume for puts that are 20 points out-of-the-money is 2,767 contracts; in contrast, the volume for puts that are 20 points in-the-money is 14 contracts. This reflects the strong demand by portfolio managers for protective puts [a phenomenon which started in late 1987 for obvious reasons].

Second, it is difficult to observe the underlying index price at the exact times that the option prices are recorded. In particular there is no guarantee that the closing index value reported is recorded at the same time as the closing transaction for each option. In particular, S&P 500 index futures are traded on the Chicago Mercantile Exchange [CME], not the CBOE, and time-stamped reported quotes may not necessarily be perfectly synchronized across the two markets. Even slight mismatches can lead to economically significant but spurious pricing anomalies.

Third, the index typically pays a dividend and the future rate of dividend payment is difficult, if not impossible, to determine. Standard and Poor's does provide daily dividend payments on the S&P500, but by nature these data are backward-looking, and there is no reason to assume that the actual dividends recorded ex-post correctly reflect the expected future dividends at the time the option is priced.

We propose to address these three problems by the following procedure. Since all option prices are recorded at the same time on each day, we require only one temporally-matched index price per day. To circumvent the unobservability of the dividend rate $\delta_{t,\tau}$, we infer the futures price $F_{t,\tau}$ for each maturity τ. By the spot- futures parity, $F_{t,\tau}$ and S_t are linked through:

$$F_{t,\tau} = S_t e^{(r_{t,\tau}-\delta_{t,\tau})\tau} \tag{5.1}$$

To derive the implied futures, we use the put-call parity relation which must hold if arbitrage opportunities are to be avoided, independently of any parametric option pricing model:[17]

$$H(S_t, X, \tau, r_{t,\tau}, \delta_{t,\tau}) \ + \ Xe^{-r_{t,\tau}\tau} \ = \ G(S_t, X, \tau, r_{t,\tau}, \delta_{t,\tau}) \ + \ F_{t,\tau}e^{-r_{t,\tau}\tau} \ . \qquad (5.2)$$

where G denotes the put price. To infer the futures price $F_{t,\tau}$ from this expression, we require reliable call and put prices—prices of actively traded options—at the *same* strike price X and time-to-expiration τ. To obtain such reliable pairs, we must use calls and puts that are closest to at-the-money [recall that in-the-money options are illiquid relative to out-the-money ones, hence any matched pair that is not at-the-money would have one potentially unreliable price]. The average daily volume for at-the-money calls and puts is 4,360 contracts and we are therefore very confident in both prices. On every day t, we do this for all available maturities τ to obtain for each maturity the implied futures price from put-call parity.

Given the derived futures price $F_{t,\tau}$, we then replace the prices of all illiquid options, i.e., in-the-money options, with the price implied by put-call parity at the relevant strike prices. Specifically, we replace the price of each in-the- money call option with $G(S_t, X, \tau, r_{t,\tau}, \delta_{t,\tau}) + F_{t,\tau}e^{-r_{t,\tau}\tau} - Xe^{-r_{t,\tau}\tau}$ where, by construction, the put with price $G(S_t, X, \tau, r_{t,\tau}, \delta_{t,\tau})$ is out-of-the-money and therefore liquid. After this procedure, all the information contained in liquid put prices has been extracted and resides in corresponding call prices via put-call parity, therefore put prices may now be discarded without any loss of reliable information.

5.2 A Nonparametric S&P 500 Index Option Pricing Formula

We use price data on every option traded during 1993, for a total of n=14,431 options after applying the filters described in the previous section. This sample size is sufficiently large to enable us to use the full nonparametric estimator as we have shown by Monte Carlo experiments. We estimated the SPD using each of the dimension reduction techniques discussed in Section 3.2.[18] The results we report below are for the semiparametric model with the kernel and bandwidth choices given in Table 3.

---

[17] Since any violation of put-call parity would give rise to a pure arbitrage opportunity, it can be expected to hold with some degree of confidence. CBOE floor traders of S&P 500 options have confirmed that put-call parity is almost never violated in practice. See also Bhattacharya (1983), Black and Scholes (1972), Chiras and Manaster (1978), Day and Lewis (1988), Galai (1977, 1978), Gould and Galai (1974), Harvey and Whaley (1992), Klemkosky and Resnick (1979, 1980), Nisbet (1992), Phillips and Smith (1980), and Rubinstein (1985).

In Figure 2 we plot the corresponding implied volatility curves for four maturities: 1, 2, 4, and 6 months. The nonparametric approach generates a strong "smile" with respect to moneyness. We follow this market's convention of quoting [and hedging] the options in terms of the futures rather than the cash index and therefore define moneyness as the ratio of strike X to futures prices $F_{t,\tau}$. The implied volatility at a fixed maturity is a decreasing nonlinear function of moneyness. Note in particular that our estimated smile is strongly asymmetric, pointing to the anecdotal evidence that out-of-money *put* prices have been consistently bid up since the crash of 1987 by investment managers looking for protection against future downward index movements. By contrast stochastic volatility models, which is the class of models most commonly used to generate smile effects, typically produce symmetric smiles [see Renault (1995)].

A further result of our approach is the changing shape of the smile as time-to-maturity increases. The one-month smile is the steepest. We find that the implied volatility curves are generally flatter for longer times-to-maturity, but we document a persistence in the smile over longer maturities that is not captured by existing stochastic volatility models. In a typical such model, mean reversion in stochastic volatility induces a rapid disappearance of the smile. Our results therefore suggest that modeling long-term memory in stochastic volatility along the lines of Harvey (1995) could be a promising approach empirically.

Note also that the curves for all maturities intersect at approximately the same level of moneyness [0.975]. In other words, options with moneyness of 0.975 are priced at about the same volatility for all maturities. At-the-money options [moneyness=1] have an implied volatility which increases slightly with maturity. The implied volatility of out-of-the-money puts [calls] decreases [increases] with maturity. This suggests that it may be misleading to focus on the term structure of at-the-money volatilities as a way of fitting the Black-Scholes model to the data for different maturities. Our nonparametric approach documents that the small differences in at-the-money implied volatilities across maturities, where all the curves are close together, understate the overall variation of implied volatilities over the full range of traded strikes. We plot in Figure 3 the three-dimensional implied volatility surface as a function of moneyness and time-to-maturity.

---

[18] Note from Table 2 that the recorded interest rate data exhibited little significant variation during 1993 and thus we could reasonably have excluded $r_{t,\tau}$ from the regressor list—that is, treating it as constant—given our sample. This would have further reduced the dimensionality of the regression. Naturally a different time period where interest rates are more volatile would require that $r_{t,\tau}$ be kept in Z for the full nonparametric model.

We report in Table 4 the nonparametric prices and deltas [with respect to the futures[19]] for a sample of calls and puts for maturities of one, two, four and six months, priced for a current futures price of 455. We give the price of every option with a delta greater or equal to 0.05 in absolute value, and the prices of the butterflies over five-point strike spreads, which by (2.11) and (2.2) give a discrete approximation to the value of the SPD at that strike level. Not surprisingly, compared to Black-Scholes prices, our prices are consistent with the features of actual market prices. Between months, we find one empirical regularity that Black-Scholes prices do not capture: the price of horizontal at-the-money straddles varies little throughout the sample. This reflects "between-month consistency" where prices for a given expiration month can vary widely, but they do so in an almost constant relation to the prices for other months. Currently we have no theoretical explanation for this empirical regularity.

5.3 S&P 500 Index Option SPDs

In Figure 4 the nonparametric SPDs for all four maturities are plotted on he same scale, emphasizing the changing shape of the distribution as maturity increases. In Figure 5 the nonparametric SPDs are overlayed with the corresponding Black-Scholes SPDs at the same maturities [the Black-Scholes log-normal SPDs are evaluated at the at-the-money implied volatility for that maturity]. Figure 5 also reports the 95% confidence intervals around each estimated SPD. The confidence interval is constructed from the asymptotic distribution theory derived in the Appendix. A simple specification test of the null hypothesis that the nonparametric family of SPDs could have been generated by the Black-Scholes model, based on the integrated squared distance between the densities, is rejected at the 5% confidence level.

---

[19] To compute deltas with respect to the stock, note that $\partial H/\partial F = (\partial H/\partial S)(\partial S/\partial F) = (\partial H/\partial S)e^{-(r-\delta)\tau}$.

Figure 6 shows the estimated nonparametric distribution of the *log-returns*, $\log(S_T/S_t)$, that is compatible with our nonparametric SPD estimator. We compute the density of log-returns by noting that:

$$\text{Prob}\left(\log\left(\frac{S_T}{S_t}\right) \leq u\right) = \text{Prob}\left(S_T \leq S_t e^u\right) = \int_0^{S_t e^u} p^*\left(S_T, S_t, \tau, r_{t,\tau}, \delta_{t,\tau}\right) dS_T \quad (5.3)$$

The density of log-returns equivalent to the SPD for prices is then:

$$\frac{\partial}{\partial u}\text{Prob}\left(\log\left(\frac{S_T}{S_t}\right) \leq u\right) = S_t e^u p^*\left(S_t e^u, S_t, \tau, r_{t,\tau}, \delta_{t,\tau}\right) \quad (5.4)$$

We compare this density to the Gaussian Black-Scholes density N($(r_{t,\tau}-\delta_{t,\tau}-\sigma^2/2)\tau$, $\sigma^2\tau$) for each maturity. Not surprisingly, the differences in the log-returns distributions are quite similar to those of the estimated SPDs for prices in Figure 5. However, computing the densities for returns allows us to illustrate the magnitude of the differences by plotting in Figure 7 the term structures of implied mean, standard deviation, skewness, and kurtosis of the SPD-generated log-return distributions along with their Black-Scholes counterparts. All the moments of the log-returns are annualized.

Figure 7 and Table 5 highlight the differences between the nonparametric and Black-Scholes SPDs. Although the nonparametric SPDs have comparable standard deviations to those obtained from the Black-Scholes formula since we estimated the Black-Scholes SPD at the actual at-the-money implied volatility, they exhibit considerably different skewness and kurtosis. Specifically, for all four maturities the nonparametric SPDs have slightly higher means, are negatively skewed, have fatter tails and the amount of skewness and kurtosis both increase with maturity.

These results point to important differences between the nonparametric SPD and the Black-Scholes SPD, which implies correspondingly important differences in the pricing implications of the two.

6 Conclusion

We have developed a nonparametric technique for estimating state-price densities based on the relation between state-price densities and option prices. Although this approach is highly data-intensive, generally requiring several thousand datapoints for a reasonable level of accuracy, it offers a promising alternative to standard parametric pricing models when parametric restrictions fail.

This trade-off between parametric restrictions and data requirements lies at the heart of the nonparametric approach—while parametric formulae are surely preferable to nonparametric ones when the underlying asset's price dynamics are well-understood, this is rarely the case in practice. Since they do not rely on restrictive parametric assumptions such as lognormality or sample-path continuity, nonparametric alternatives are robust to the specification errors that plague parametric models. A nonparametric approach is particularly valuable in such applications since the typical parametric restrictions have been shown to fail, sometimes dramatically. For example, Figure 7 shows that the primary failure of the Black-Scholes model in pricing S&P 500 index options is its inability to account for the skewness and kurtosis apparent in the nonparametric estimates of the returns distribution, and even more so for longer-maturity options. Parametric extensions of the Black-Scholes model should focus on capturing these empirical facts. The amount of persistence in the smile is such that parametric models incorporating long-term memory in stochastic volatility may be the most promising.

Also, by their very nature nonparametric methods are adaptive, responding to structural shifts in the data-generating processes in ways that parametric models do not. And finally, they are flexible enough to encompass a wide range of derivative securities and fundamental asset price dynamics, yet relatively simple and computationally efficient to implement.

Appendix

In this Appendix we describe in more detail our procedure for selecting the kernel functions and the bandwidths in our option-pricing context. We also discuss the rate of convergence of our kernel estimator of the call pricing function to the true function, its partial derivative with respect to the asset price to the true option delta and its second derivative with respect to the strike price to the true SPD.

Kernel and Bandwidth Selection

We have to choose both the univariate kernel function k and the bandwidth parameters in h for each regressor. We will differentiate twice the right-hand-side of (3.4) with respect to the strike price X to obtain an estimate of $\partial^2 H/\partial X^2$, and once with respect to S_t to estimate the option delta. We therefore require that k_X [k_S] be at least twice [once] continuously-differentiable.

Four elements determine the choice of the kernel and bandwidth: the sample size n, i.e., the number of options used to construct the estimator $\hat{H}(\cdot)$, the total number d of regressors included in the nonparametric regression, the number p_j of existing continuous partial derivatives of the true option pricing function $H(\cdot)$ with respect to the j-th regressor Z_j and finally the order m_j of the partial derivative with respect to the j-th regressor that we wish to estimate [with the convention that $m_j = 0$ when no partial differentiation is required with respect to Z_j].

We naturally assume that the call pricing function $H(\cdot)$ to be estimated is sufficiently smooth, i.e., $p_j \geq m_j$ for all $j = 1, \ldots, d$. In our problem, the highest derivative that we will estimate is the second partial derivatives of the call price with respect to the strike price. We assume that the function $H(\cdot)$ admits four continuous derivatives with respect to each of its regressors, that is $p_j = p = 4$ for all $j = 1, \ldots, d$.[20]

---

[20] It is possible to make primitive assumptions on the data-generating process which imply the necessary smoothness of the true call-pricing function. In particular, if we assumed that the underlying asset price followed a stochastic differential equation with diffusion function σ^2 which admitted at least p continuous derivatives, $p \geq 2$, then the call pricing function would also admit at least p continuous derivatives. This follows by writing the call pricing function as the solution of the generalized Black-Scholes partial differential equation derived from using the standard dynamic replicating strategy [see Merton (1973)]. It is a parabolic

Define the order s of a kernel function k as the even integer that satisfies the following relation: $\int z^l \mathrm{k}(z)du = 0$ for $l = 1, \ldots, s-1$, and $\int |z|^s \mathrm{k}(z)du < \infty$. For example, the popular Gaussian kernel:

$$\mathrm{k}^{(2)}(z) \;=\; \frac{1}{\sqrt{2\pi}} e^{-z^2/2} \tag{A.1}$$

is of order $s = 2$, while the following kernel is of order $s = 4$:

$$\mathrm{k}^{(4)}(z) \;=\; \frac{3}{\sqrt{8\pi}} \left(1 - \frac{z^2}{3}\right) e^{-z^2/2} \,. \tag{A.2}$$

Using higher order kernels has the effect of accelerating the speed of convergence of the estimate to the true function as the sample size increases, in a mean-squared sense. We set the order of the kernel for each regressor j to be $s_j = p - m_j$ [if even, $s_j = p - m_j - 1$ otherwise].[21] However higher order kernels can be cumbersome to use in practice- -they are no longer uniformly positive for example. Experience suggests that it is desirable to limit the choice to kernels of orders no larger than $s = 4$.

Therefore, to estimate the call pricing function where $m_j = 0$ for each regressor we set in the full nonparametric model $\mathrm{k}_X = \mathrm{k}_S = \mathrm{k}_\tau = \mathrm{k}_r = \mathrm{k}_\delta = \mathrm{k}^{(4)}$, and in the semiparametric model (3.5), $\mathrm{k}_X = \mathrm{k}_S = \mathrm{k}_\tau = \mathrm{k}^{(4)}$. To estimate the option delta where $m_j = 0$ for each regressor except S_t for which $m_j = 1$: in the full nonparametric model, we set $\mathrm{k}_X = \mathrm{k}_\tau = \mathrm{k}_r = \mathrm{k}_\delta = \mathrm{k}^{(4)}$ and $\mathrm{k}_S = \mathrm{k}^{(2)}$, and in the semiparametric model, $\mathrm{k}_X = \mathrm{k}_\tau = \mathrm{k}^{(4)}$ and $\mathrm{k}_S = \mathrm{k}^{(2)}$. Finally, to estimate the SPD where $m_j = 0$ for each regressor except X for which $m_j = 2$: in the full nonparametric model, we set $\mathrm{k}_S = \mathrm{k}_\tau = \mathrm{k}_r = \mathrm{k}_\delta = \mathrm{k}^{(4)}$ and $\mathrm{k}_X = \mathrm{k}^{(2)}$, and in the semiparametric model, $\mathrm{k}_S = \mathrm{k}_\tau = \mathrm{k}^{(4)}$ and $\mathrm{k}_X = \mathrm{k}^{(2)}$.

Monte Carlo evidence in Section 4 shows that for the choices of kernel functions above,

---

differential equation satisfying all the hypotheses in Friedman (1964, Chapter I) whose solution is then known to have at least p derivatives.

[21] When estimating the m_j-th partial derivative of $H(\cdot)$ with respect to the j-th regressor, we obtain a bias term of order $h_j^{p-m_j}$ and a variance term of order $1/(nh_j^{(2m_j+d)})$. We are minimizing the mean-squared error of the estimate, which consists of the sum of the squared bias and variance terms. Since the bandwidth h_j goes to zero as the sample size n grows, the larger the order of the kernel, the lower the bias term [the curve fit improves] but the larger the variance term [the curve estimate becomes noisier]. Our choice of the bandwidth given in (A.3) optimally balances these two effects.

$\hat{H}(\cdot)$ is very accurate for the typically large sample size n considered in our empirical application.

Provided that the correct order of the kernel for each regressor is determined, the choice of the function k over a large set of reasonable candidates has little influence on the estimator of $\hat{H}(\cdot)$. However the bandwidth choice is crucial. The larger the bandwidth, the smoother the curve estimate as a function of the regressors. Deciding how much to smooth, i.e., how fast to decrease the bandwidth to zero as a function of the sample size, is the essential component of nonparametric regression analysis.

For each of the d regressors in Z, we set the corresponding bandwidth parameter h_j according to the relation:

$$h_j \quad = \quad c_j \sigma(Z_j) n^{-1/(d+2p)} \tag{A.3}$$

where $\sigma(Z_j)$ is the unconditional standard deviation of the regressor Z_j, $j = 1, \ldots, d$ and c_j a fixed constant. This bandwidth choice is such that our estimator $\hat{H}(\cdot)$ achieves the optimal rate of convergence in the mean-squared sense among all possible nonparametric estimators of $H(\cdot)$:

$$n^{(p-m)/(d+2p)} \tag{A.4}$$

where $m \equiv \sum_{j=1}^{d} m_j$.

Rules to choose the constant c_j in (A.3) can be found in Härdle (1990, pp. 161-162). c_j depends on the choice of the kernel and the function to be estimated. c_j is typically of the order of one and small deviations from the exact value have no large effects. In practice, we select the constant c_j by *cross-validation* [see Härdle (1990) for example], a technique which ensures that we minimize the mean-squared error of our estimator $\hat{H}(\cdot)$.

Note from (A.4) that the lower the dimensionality d of the regression function, i.e., the smaller the number of regressors in Z, the faster the convergence of the estimator $\hat{H}(Z) = \hat{\mathrm{E}}[H|Z]$ and its derivatives to the true function [see (A.4)]. This is the technical motivation for our proposed approaches to dimension direction in Section 3.2.

Note also from (A.4) that higher-order derivatives converge at a slower speeds, hence the SPD estimator [for which $m = 2$] converges slower than the delta estimator [for which $m = 1$], which in turn converges slower than the price estimator [for which $m = 0$]. One additional order of differentiation slows down the rate of convergence as much as two additional regressors, i.e., the "curse of differentiation" [the decrease in rate of convergence as m increases] is twice as damning as the "curse of dimensionality" [the decrease in rate of convergence as d increases]. As a theoretical matter, we can still get arbitrarily close to the parametric rate of convergence $n^{1/2}$ if the call pricing function has enough continuous derivatives, and we use a kernel of sufficiently high order $p - m$: as p increases, the rate of convergence $n^{(p-m)/(2p+d)}$ converges to the parametric rate $n^{1/2}$.

Asymptotic Distributions

Using the kernel and bandwidth selections just described, we obtain the following from the general results in Aït-Sahalia (1993), where precise regularity conditions are stated:

$$n^{p/(2p+4)}\Big[\,\hat{H}(S_t,X,\tau,r_{t,\tau},\delta_{t,\tau}) - H(S_t,X,\tau,r_{t,\tau},\delta_{t,\tau})\,\Big] \quad \overset{a}{\sim} \quad N(0,\sigma_c^2) \tag{A.5}$$

$$n^{(p-1)/(2p+4)}\Big[\,\hat{\Delta}(S_t,X,\tau,r_{t,\tau},\delta_{t,\tau}) - \Delta(S_t,X,\tau,r_{t,\tau},\delta_{t,\tau})\,\Big] \quad \overset{a}{\sim} \quad N(0,\sigma_\Delta^2) \tag{A.6}$$

$$n^{(p-2)/(2p+4)}\Big[\,\hat{p}^*(S_t,S_T,\tau,r_{t,\tau},\delta_{t,\tau}) - p^*(S_t,S_T,\tau,r_{t,\tau},\delta_{t,\tau})\,\Big] \quad \overset{a}{\sim} \quad N(0,\sigma_p^2) \tag{A.7}$$

where $\epsilon \equiv H(S_t,X,\tau,r_{t,\tau},\delta_{t,\tau}) - \mathrm{E}[H|S_t,X,\tau,r_{t,\tau},\delta_{t,\tau}]$ and

$$\sigma_c^2 \quad \equiv \quad \frac{\mathrm{E}[\epsilon^2|S_t,X,\tau,r_{t,\tau},\delta_{t,\tau}]\Big(\int_{-\infty}^{\infty} k^2(w)dw\Big)^4}{\pi(S_t,X,\tau,r_{t,\tau},\delta_{t,\tau})} \tag{A.8}$$

$$\sigma_\Delta^2 \quad \equiv \quad \frac{\mathrm{E}[\epsilon^2|S_t,X,\tau,r_{t,\tau},\delta_{t,\tau}]\int_{-\infty}^{\infty}[k'(w)]^2dw\Big(\int_{-\infty}^{\infty} k^2(w)dw\Big)^3}{\pi(S_t,X,\tau,r_{t,\tau},\delta_{t,\tau})} \tag{A.9}$$

$$\sigma_p^2 \quad \equiv \quad \frac{\mathrm{E}[\epsilon^2|S_t,X,\tau,r_{t,\tau},\delta_{t,\tau}]\int_{-\infty}^{\infty}[k''(w)]^2dw\Big(\int_{-\infty}^{\infty} k^2(w)dw\Big)^3}{\pi(S_t,X,\tau,r_{t,\tau},\delta_{t,\tau})} \tag{A.10}$$

and $\pi(S_t, X, \tau, r_{t,\tau}, \delta_{t,\tau})$ is the joint density function of the variables S_t, X, τ, $r_{t,\tau}$ and $\delta_{t,\tau}$. These asymptotic distributions give immediately the relevant confidence intervals for the nonparametric estimators of the call pricing function, delta, and SPD respectively.

References

Aït-Sahalia, Y., 1992, "Nonparametric Pricing of Interest Rate Derivative Securities", to appear in *Econometrica.*

Aït-Sahalia, Y., 1993, "The Delta Method for Nonparametric Kernel Functionals", working paper, University of Chicago.

Aït-Sahalia, Y., 1994, "Testing Continuous-Time Models of the Spot Interest Rate", to appear in *Review of Financial Studies.*

Arrow, K., 1964, "The Role of Securities in the Optimal Allocation of Risk Bearing", *Review of Economic Studies* 31, 91–96.

Banz, R. and M. Miller, 1978, "Prices for State- Contingent Claims: Some Estimates and Applications", *Journal of Business* 51, 653–672.

Bates, D.S., 1995, "Post-'87 Crash Fears in S&P 500 Futures Options", working paper, The Wharton School.

Barlow, R.E., D.J. Bartholomew, J.M. Bremmer and H.D. Brunk, 1972, *Statistical Inference Under Order Restrictions.* New York: Wiley.

Bick, A., 1990, "On Viable Diffusion Price Processes of the Market Portfolio", *Journal of Finance* 45, 673–689.

Bhattacharya, M., 1983, "Transactions Data Tests of Efficiency of the Chicago Board Options Exchange", *Journal of Financial Economics* 12, 161–185.

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

Black, F. and M. Scholes, 1972, "The Valuation of Option Contracts and a Test of Market Efficiency", *Journal of Finance* 27, 399–418.

Boudoukh, J., Richardson, M., Stanton, R. and R. Whitelaw, 1995, "Pricing Mortgage-Backed Securities in a Multifactor Interest Rate Environment: A Multivariate Density Estimation Approach", working paper, Stern School of Business, New York University.

Boyle, P. and D. Emanuel, 1980, "Discretely Adjusted Option Hedges", *Journal of Financial Economics* 8, 259–82.

Breeden, D. and R. Litzenberger, 1978, "Prices of State-Contingent Claims Implicit in Option Prices", *Journal of Business* 51, 621–651.

Chiras, D. and S. Manaster, 1978, "The Informational Content of Option Prices and a Test of Market Efficiency", *Journal of Financial Economics* 6, 213–234.

Constantinides, G., 1982, "Intertemporal Asset Pricing with Heterogeneous Consumers and without Demand Aggregation", *Journal of Business* 55, 253–268.

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

Cox, J., Ross, S. and M. Rubinstein, 1979, "Option Pricing: A Simplified Approach", *Journal of Financial Economics* 7, 229–263.

Cox, J. and M. Rubinstein, 1985, *Options Markets.* Englewood Cliffs, New Jersey: Prentice-Hall.

Day, T. and C. Lewis, 1988, "The Behavior of Volatility Implicit in the Prices of Stock Index Options", *Journal of Financial Economics* 22, 103–122.

Debreu, G., 1959, *Theory of Value.* New York: John Wiley and Sons.

Derman, E. and I. Kani, 1994, "Riding on the Smile", *RISK* 7 (February), 32–39.

Dothan, M., 1990, *Prices in Financial Markets.* New York: Oxford University Press.

Duffie, D., 1992, *Dynamic Asset Pricing Theory.* Princeton, NJ: Princeton University Press.

Duffie, D. and C. Huang, 1985, "Implementing Arrow-Debreu Equilibria by Continuous Trading of Few Long-Lived Securities", *Econometrica* 53, 1337–1356.

Dupire, B., 1994, "Pricing with a Smile", *RISK* 7 (January), 18–20.

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

Friedman, A., 1964, *Partial Differential Equations of Parabolic Type.* Englewood Cliffs, NJ: Prentice-Hall.

Galai, D., 1977, "Tests of Market Efficiency and the Chicago Board Options Exchange", *Journal of Business* 50, 167–197.

Galai, D., 1978, "Empirical Tests of Boundary Conditions for CBOE Options", *Journal of Financial Economics* 6, 187–211.

Garman, M., 1978, "The Pricing of Supershares", *Journal of Financial Economics* 6, 3–10.

Goldenberger, D.H., 1991, "A Unified Method for Pricing Options on Diffusion Processes", *Journal of Financial Economics* 29, 3-34.

Gould, J. and D. Galai, 1974, "Transaction Costs and the Relationship Between Put and Call Prices", *Journal of Financial Economics* 1, 105–129.

Grundy, B., 1991, "Option Prices and the Underlying Asset's Return Distribution", *Journal of Finance* 46, 1045–1069.

Hansen, L.P. and S.F. Richard, 1987, "The Role of Conditioning Information in Deducing Testable Restrictions Implied by Dynamic Asset Pricing Models", *Econometrica* 55, 3, 587-613.

Hansen, L.P. and R. Jagannathan, 1991, "Implications of Security Market Data for Models of Dynamic Economies", *Journal of Political Economy* 99, 225–262.

Hakansson, N., 1976, "Purchasing Power Funds: A New Kind of Financial Intermediary", *Financial Analysts Journal* 32, 49–59.

Hakansson, N., 1977, "The Superfund: Efficient Paths Towards Efficient Capital Markets in Large and Small Countries", in *Financial Decision Making Under Uncertainty*, edited by H. Levy and M. Sarnat, New York: Academic Press.

Härdle, W., 1990, *Applied Nonparametric Regression.* Cambridge, UK: Cambridge University Press.

Harrison, M. and D. Kreps, 1979, "Martingales and Arbitrage in Multiperiod Securities Markets", *Journal of Economic Theory* 20, 381–408.

Harvey, A., 1995, "Long Memory in Stochastic Volatility", working paper, London School of Economics.

Harvey, C. and R. Whaley, 1992, "Market Volatility Prediction and the Efficiency of the S&P 100 Index Option Market", 43–73.

He, H. and H. Leland, 1993, "On Equilibrium Asset Price Processes", *Review of Financial Studies* 6, 593–617.

Heston, S.L., 1993, "A Closed-form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options", *Review of Financial Studies* 6, 327-343.

Huang, C. and R. Litzenberger, 1988, *Foundations For Financial Economics.* New York: Elsevier Publishing Company.

Hull, J., 1993, *Options, Futures, and Other Derivative Securities*, 2nd edition. Englewood Cliffs, New Jersey: Prentice-Hall.

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

Hutchinson, J., Lo, A. and T. Poggio, 1994, "A Nonparametric Approach to the Pricing and Hedging of Derivative Securities Via Learning Networks", *Journal of Finance* 49, 851–889.

Jackwerth, J.C. and M. Rubinstein, 1995, "Recovering Probability Distributions from Contemporary Security Prices', working paper, Haas School of Business, University of California at Berkeley.

Jarrow, R. and A. Rudd, 1982, "Approximate Option Valuation for Arbitrary Stochastic Processes", *Journal of Financial Economics* 10, 347-369.

Klemkosky, R. and B. Resnick, 1979, "Put-Call Parity and Market Efficiency", *Journal of Finance* 34, 1141–1155.

Klemkosky, R. and B. Resnick, 1980, "An Ex Ante Analysis of Put-Call Parity", *Journal of Financial Economics* 8, 363–378.

Leland, H., 1985, "Option Pricing and Replication with Transactions Costs", *Journal of Finance* 40, 1283–1301.

Lo, A., 1986, "Statistical Tests of Contingent Claims Asset-Pricing Models: A New Methodology", *Journal of Financial Economics* 17, 143–173.

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

Lo, A. and A. Craig MacKinlay, 1988, "Stock Market Prices Do Not Follow Random Walks: Evidence from a Simple Specification Test", *Review of Financial Studies* 1, 41-66.

Longstaff, F., 1992, "An Empirical Examination of the Risk-Neutral Valuation Model", working paper, College of Business, Ohio State University, and the Anderson Graduate School of Management, UCLA.

Longstaff, F., 1994, "Option Pricing and the Martingale Restriction", to appear in *Review of Financial Studies.*

Lucas, R., 1978, "Asset Prices in an Exchange Economy", *Econometrica* 46, 1429–1446.

Madan, D.B. and F. Milne, 1994, "Contingent Claims Valued and Hedged by Pricing and Investing in a Basis", *Mathematical Finance* 4, 223-245.

Mammen, E., 1991, "Estimating a Smooth Monotone Regression Function", *The Annals of Statistics* 19, 724-740.

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

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

Merton, R., 1976, "Option Pricing When Underlying Stock Returns are Discontinuous", *Journal of Financial Economics* 3, 125–144.

Merton, R., 1992, *Continuous-Time Finance*, Revised Edition. Oxford, UK: Basil Blackwell.

Nelson, D.B. and K. Ramaswamy, 1990, "Simple Binomial Processes as Diffusion Approximations in Financial Models", *Review of Financial Studies* 3, 393-430.

Nisbet, M., 1992, "Put-Call Parity Theory and An Empirical Test of the Efficiency of the London Trade Option Market", *Journal of Banking and Finance* 16, 381–404.

Rady, S., 1994, "State Prices Implicit in Valuation Formulae for Derivative Securities: A Martingale Approach", Discussion Paper No. 181, LSE Financial Markets Group, London, UK.

Renault, E., 1995, "Econometric Models of Option Pricing Errors", working paper, Université de Toulouse; to appear in the collected invited papers at the 7th World Congress of the Econometric Society.

Ross, S., 1976, "Options and Efficiency", *Quarterly Journal of Economics* 90, 75–89.

Rubinstein, M., 1976, "The Valuation of Uncertain Income Streams and the Pricing of Options", *Bell Journal of Economics*, 407-425.

Rubinstein, M., 1985, "Nonparametric Tests of Alternative Option Pricing Models Using All Reported Trades and Quotes on the 30 Most Active CBOE Option Classes from August 23, 1976 Through August 31, 1978", *Journal of Finance* 40, 455–480.

Rubinstein, M., 1994, "Implied Binomial Trees", *Journal of Finance* 49, 771–818.

Shimko, D., 1993, "Bounds of Probability", *RISK* 6, 33–37.

Stutzer, M., 1995, "A Simple Nonparametric Approach to Derivative Security Valuation", working paper, Carlson School of Management, University of Minnesota.

Wand, M.P. and M.C. Jones, 1995, *Kernel Smoothing*, London, UK: Chapman and Hall.

Table 1

Monte Carlo simulation of tracking errors for delta-hedged arbitrage portfolios using a nonparametric kernel option-pricing formula and the Black-Scholes formula, under the Black-Scholes assumptions. At inception, each simulated arbitrage portfolio contains a single 252-day call option, and stock and bond holdings that yield a zero net investment for the portfolio. Portfolios are rebalanced every 5 days to delta-hedge the option, and the terminal values V_T of the portfolios are used to measure the tracking error of the delta-hedging strategies over the course of 252-day sample paths. The values of the initial index level, the interest rate, and the index return mean and standard deviation are fixed at 455, 3%, 7.95%, and 10.28% respectively. Each strike price column reports the results of 5,000 independent replications. Standard errors are reported in parentheses.

Tracking Error Measure	Strike = 415		Strike = 455		Strike = 495	
	Kernel	BS	Kernel	BS	Kernel	BS
$e^{-rT}\mathrm{E}[V_T]$	0.045 (0.038)	0.037 (0.034)	0.051 (0.052)	0.027 (0.045)	0.135 (0.085)	0.040 (0.070)
$e^{-rT}\mathrm{E}[\lvert V_T\rvert]$	0.620 (0.030)	0.662 (0.027)	1.442 (0.035)	1.475 (0.030)	1.628 (0.059)	1.651 (0.047)
$100 \times e^{-rT}\mathrm{E}[V_T\ ]/C_{\mathrm{BS},0}$	0.089 (0.076)	0.073 (0.067)	0.224 (0.229)	0.120 (0.198)	1.856 (1.165)	0.547 (0.969)
$100 \times e^{-rT}\mathrm{E}[\lvert V_T\rvert]/C_{\mathrm{BS},0}$	1.219 (0.060)	1.300 (0.053)	6.368 (0.153)	6.515 (0.134)	22.388 (0.802)	22.702 (0.651)

Table 2

Summary statistics for the sample of all traded CBOE daily call and put option prices on the S&P 500 index from January 4, 1993 to December 31, 1993 (14,431 observations). "Implied σ" denotes the implied volatility of the option, and "Implied ATM σ" denotes the implied volatility of at-the-money options. During this period, the sample daily mean and standard deviation of continuously-compounded returns of the S&P 500 index was 7.95% (annualized with a 365-day year) and 10.28% (annualized with a 252-day year), respectively.

Variable	Mean	S.D.	Min	Percentiles					Max
				5%	10%	50%	90%	95%	
Call Price H ($)	24.23	25.41	0.13	0.32	0.76	16.68	60.08	74.50	121.93
Put Price G ($)	9.75	12.57	0.13	0.30	0.54	4.73	25.50	33.32	102.08
Implied σ (%)	11.36	3.29	5.07	7.44	7.82	10.71	15.66	17.41	36.43
Implied ATM σ (%)	9.37	0.86	6.10	7.96	8.28	9.36	10.38	10.65	16.47
τ (Days)	86.64	72.32	1.00	11.00	21.00	66.00	196.00	259.00	350.00
X (Index Points)	440.80	33.02	350.00	390.00	400.00	440.00	480.00	490.00	550.00
F (Index Points)	455.42	10.26	428.70	436.13	441.33	457.82	467.46	469.11	474.44
r (%)	3.07	0.08	2.85	2.96	2.97	3.08	3.18	3.19	3.21

Table 3

Optimal bandwidth selection for SPD estimation with kernels k_S, k_X, and k_τ according to the relation $h_j = c_j \sigma_j n^{-1/[2p+d]}$ for regressor j, where n is the sample size, p is the number of continuous derivatives of the function to be estimated, d is the number of regressors, σ_j is the unconditional variance of the regressor, and c_j is a constant which depends on the particular regressor and is determined by cross-validation. m_j is the order of the partial derivative with respect to the regressor that we wish to estimate, and s_j is the order of the corresponding kernel [see the Appendix for further details]. As a measure of the goodness-of-fit of our nonparametric kernel regression, we compute the coefficient of determination $R^2 = 0.86$.

Kernel	s_j	p	m_j	d	c_j	σ_j	h_j
$\mathrm{k}_S = \mathrm{k}^{(4)}$	4	4	0	3	2.040	10.265	8.767
$\mathrm{k}_X = \mathrm{k}^{(2)}$	2	4	2	3	1.260	33.018	17.418
$\mathrm{k}_\tau = \mathrm{k}^{(4)}$	4	4	0	3	0.1014	72.324	3.071

Table 4

Estimated nonparametric call option, put option, and butterfly-spread prices on the S&P 500 index for one-, two-, four- and six-month maturities and strike prices with deltas greater than or equal to 0.05 in absolute value, priced for a current index value of 455.00. The nonparametric kernel estimator is based on a sample of 14,431 CBOE daily call and put option prices on the S&P 500 index from January 4, 1993 to December 31, 1993 [see Table 2 for further details]. Prices for horizontal at-the-money straddles [strike price 455.00] are: $3.831 for the 42-day/21-day straddle, $6.625 for the 84-day/42-day straddle, and $5.506 for the 126-day/84-day straddle. Option deltas are computed with respect to the futures price, not the spot price.

Strike Price	Call Price	Call Delta	Put Price	Put Delta	Implied Volatility	Butterfly Price
Futures: $455.4843, Interest Rate: 2.9%, Time-to-Maturity: 21 Days						
435	21.21	0.94	0.76	−0.05	12.48	0.17
440	16.68	0.90	1.22	−0.10	11.75	0.28
445	12.42	0.83	1.95	−0.17	11.07	0.44
450	8.61	0.72	3.13	−0.28	10.44	0.63
455	5.42	0.57	4.94	−0.43	9.89	0.79
460	3.03	0.40	7.54	−0.60	9.43	0.83
465	1.46	0.24	10.96	−0.76	9.05	0.70
470	0.60	0.12	15.09	−0.88	8.74	0.47
Futures: $456.0209, Interest Rate: 3.00%, Time-to-Maturity: 42 Days						
430	27.18	0.93	1.25	−0.07	12.11	0.14
435	22.70	0.90	1.75	−0.10	11.56	0.20
440	18.40	0.85	2.44	−0.14	11.02	0.28
445	14.39	0.79	3.41	−0.21	10.50	0.38
450	10.75	0.70	4.75	−0.30	10.02	0.49
455	7.61	0.58	6.59	−0.41	9.58	0.58
460	5.04	0.46	9.01	−0.54	9.21	0.63
465	3.11	0.33	12.05	−0.67	8.89	0.60
470	1.76	0.21	15.70	−0.78	8.63	0.50
475	0.92	0.12	19.83	−0.87	8.42	0.36
480	0.44	0.06	24.34	−0.93	8.25	0.23

Table 4 (continued)

Strike Price	Call Price	Call Delta	Put Price	Put Delta	Implied Volatility	Butterfly Price
Futures: $457.0963, Interest Rate: 3.05%, Time-to-Maturity: 84 Days						
420	38.73	0.91	1.89	−0.08	12.56	0.10
425	34.27	0.89	2.40	−0.11	12.18	0.13
430	29.94	0.85	3.03	−0.14	11.80	0.16
435	25.76	0.81	3.82	−0.18	11.42	0.19
440	21.78	0.77	4.81	−0.23	11.04	0.24
445	18.04	0.71	6.02	−0.28	10.67	0.28
450	14.57	0.65	7.53	−0.35	10.30	0.34
455	11.45	0.57	9.37	−0.42	9.95	0.39
460	8.71	0.49	11.60	−0.50	9.62	0.42
465	6.40	0.40	14.25	−0.59	9.32	0.44
470	4.52	0.32	17.34	−0.67	9.06	0.42
475	3.07	0.24	20.85	−0.75	8.82	0.38
480	1.99	0.17	24.74	−0.82	8.62	0.32
485	1.24	0.11	28.95	−0.88	8.44	0.25
490	0.74	0.07	33.41	−0.92	8.30	0.18
Futures: $458.2245, Interest Rate: 3.10%, Time-to-Maturity: 126 Days						
420	41.12	0.93	3.31	−0.06	12.53	0.10
425	36.85	0.91	3.98	−0.08	12.20	0.13
430	32.70	0.88	4.77	−0.11	11.87	0.15
435	28.70	0.84	5.72	−0.15	11.55	0.18
440	24.88	0.80	6.85	−0.19	11.22	0.21
445	21.26	0.75	8.18	−0.24	10.91	0.23
450	17.88	0.69	9.74	−0.30	10.59	0.26
455	14.76	0.63	11.57	−0.36	10.28	0.29
460	11.92	0.57	13.68	−0.42	9.98	0.32
465	9.41	0.50	16.11	−0.49	9.69	0.35
470	7.24	0.44	18.89	−0.55	9.42	0.36
475	5.44	0.37	22.03	−0.62	9.18	0.35
480	3.98	0.30	25.52	−0.69	8.97	0.32
485	2.84	0.24	29.33	−0.75	8.80	0.28
490	1.98	0.18	33.42	−0.81	8.65	0.23
495	1.35	0.14	37.73	−0.85	8.53	0.18
500	0.90	0.10	42.23	−0.89	8.42	0.14
505	0.58	0.07	46.86	−0.92	8.34	0.10

Table 5

Moments of nonparametric and Black-Scholes log-returns density estimators based on a sample of 14,431 CBOE daily call and put option prices on the S&P 500 index from January 4, 1993 to December 31, 1993 [see Table 2 for further details].

SPD Estimator	Mean	Standard Deviation	Skewness	Kurtosis
Time-to-Maturity: 21 Days				
Nonparametric	0.0233	0.0991	−0.1972	0.0741
Black-Scholes	0.0136	0.0991	0.0000	0.0000
Time-to-Maturity: 42 Days				
Nonparametric	0.0183	0.0981	−0.3381	0.2154
Black-Scholes	0.0149	0.0952	0.0000	0.0000
Time-to-Maturity: 84 Days				
Nonparametric	0.0179	0.0995	−0.4337	0.2170
Black-Scholes	0.0152	0.0981	0.0000	0.0000
Time-to-Maturity: 126 Days				
Nonparametric	0.0225	0.0993	−0.5051	0.2798
Black-Scholes	0.0154	0.1001	0.0000	0.0000

Figure 1a
Black-Scholes Option Price

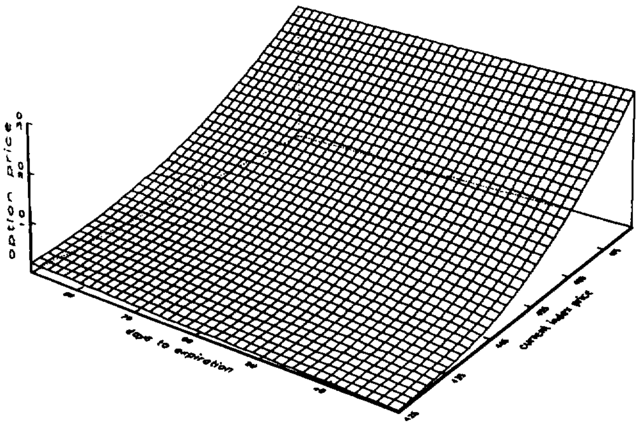


Figure 1b
Option Price % Difference

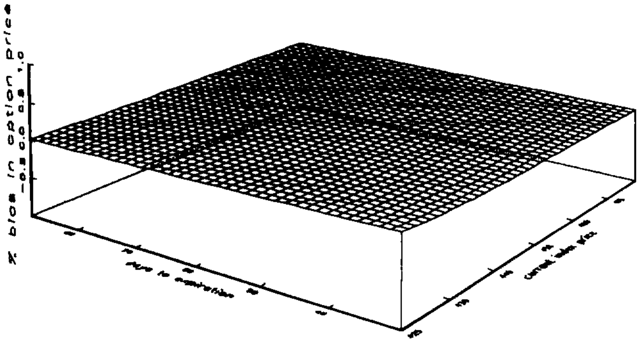


Figure 1c
Black-Scholes Option Delta

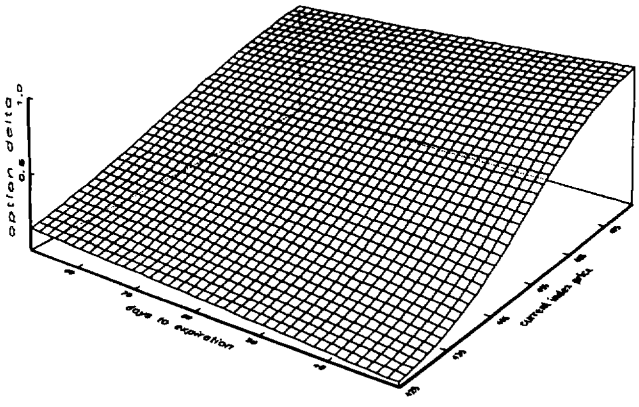


Figure 1d
Option Delta % Difference

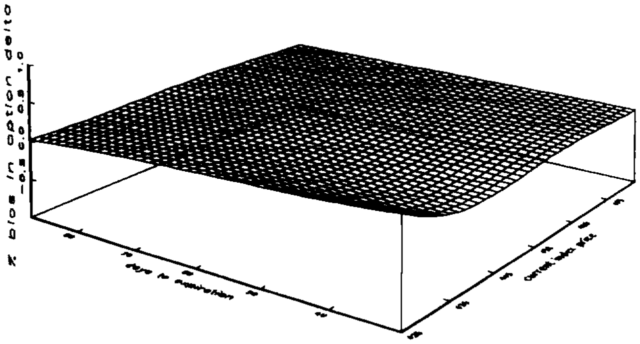


Figure 1e
Black-Scholes State-Price Density

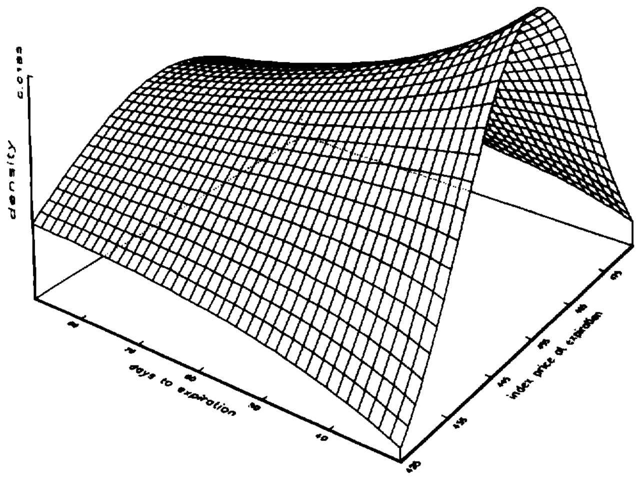


Figure 1f
State-Price Density % Difference

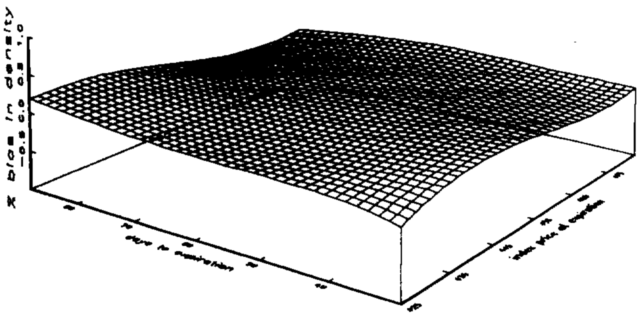


Figure 2
SPX Estimated Implied Volatilities

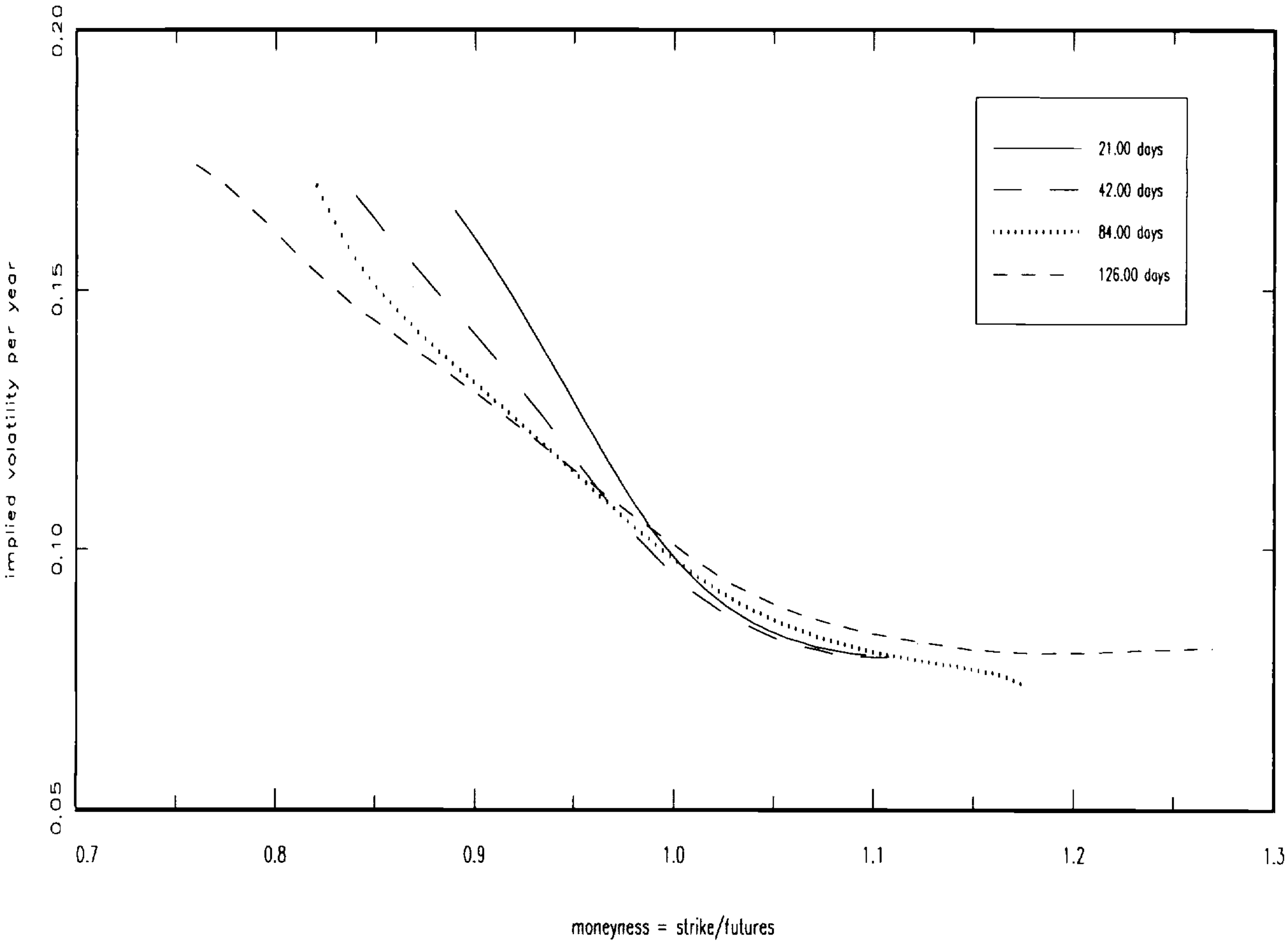


Figure 3

SPX Estimated Implied Volatility Surface

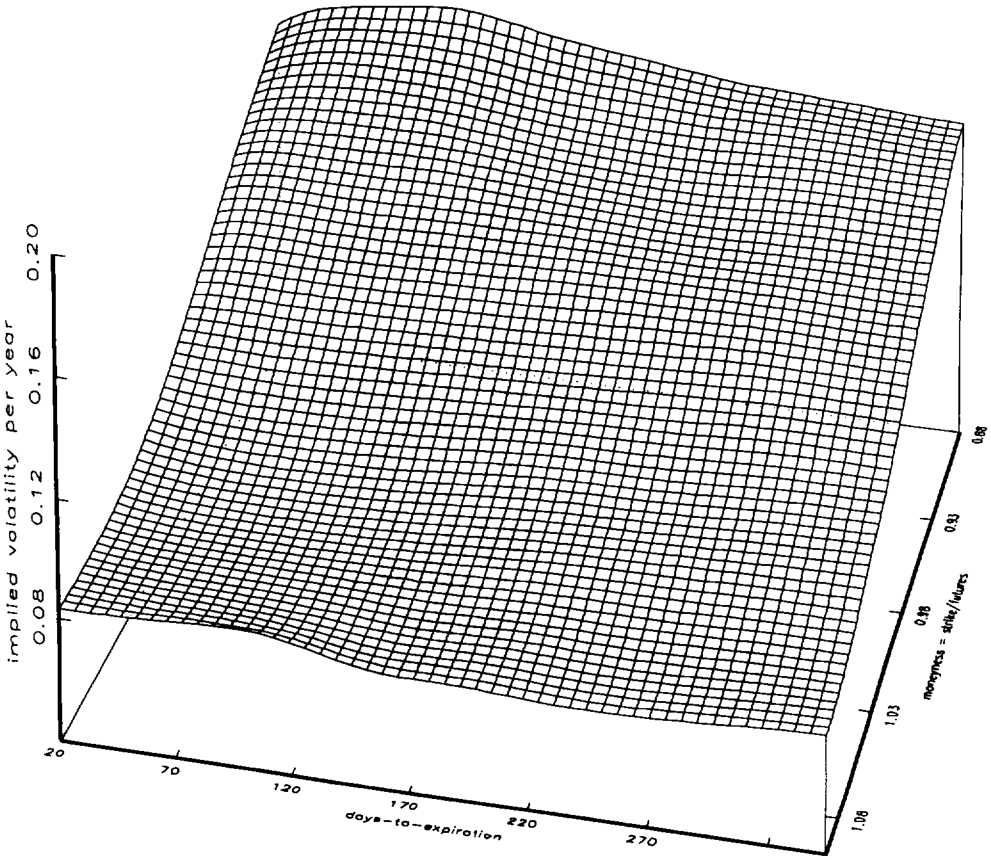


Figure 4
SPX Estimated Nonparametric State-Price Densities

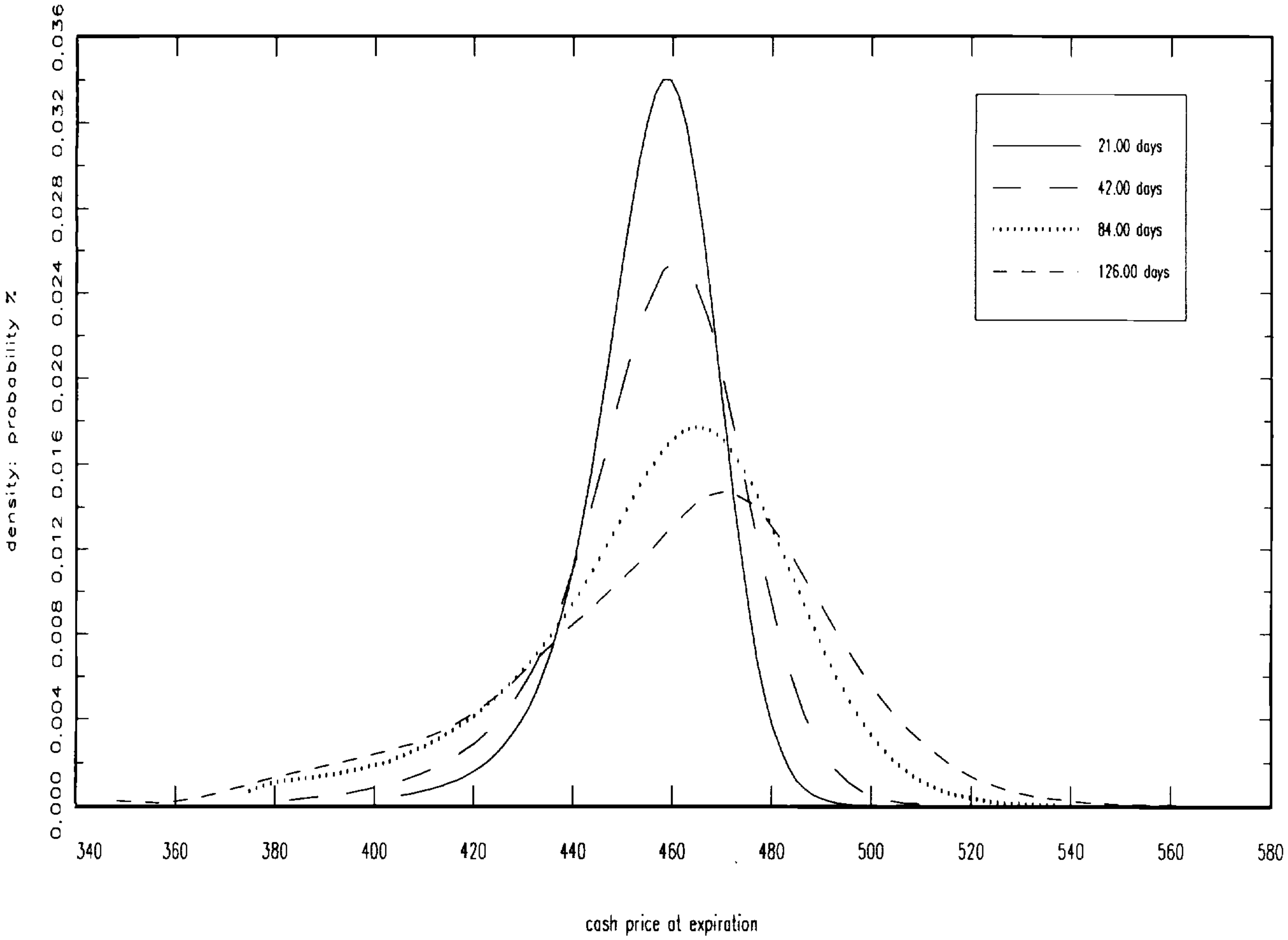


Figure 5
SPX Estimated State-Price Densities
Maturity =21.00 days

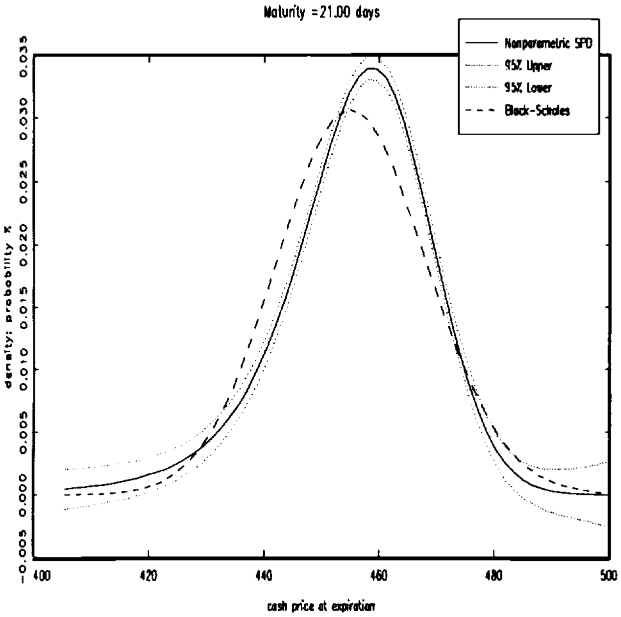


Figure 5
SPX Estimated State-Price Densities
Maturity =42.00 days

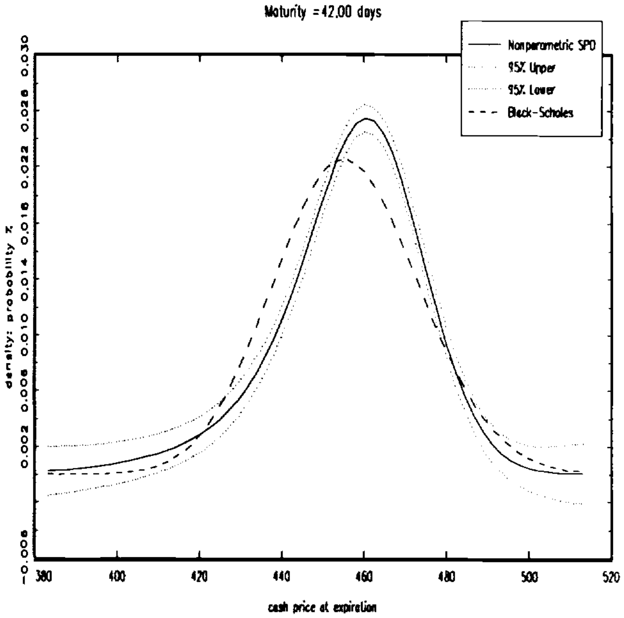


Figure 5
SPX Estimated State-Price Densities
Maturity =84.00 days

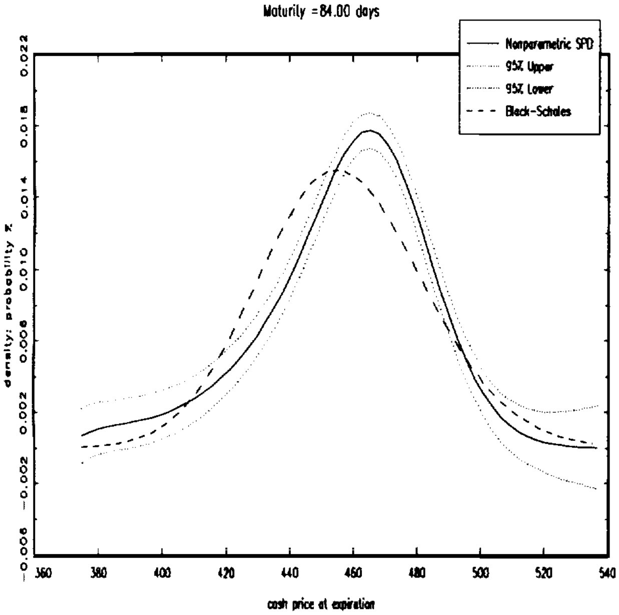


Figure 5
SPX Estimated State-Price Densities
Maturity =126.00 days

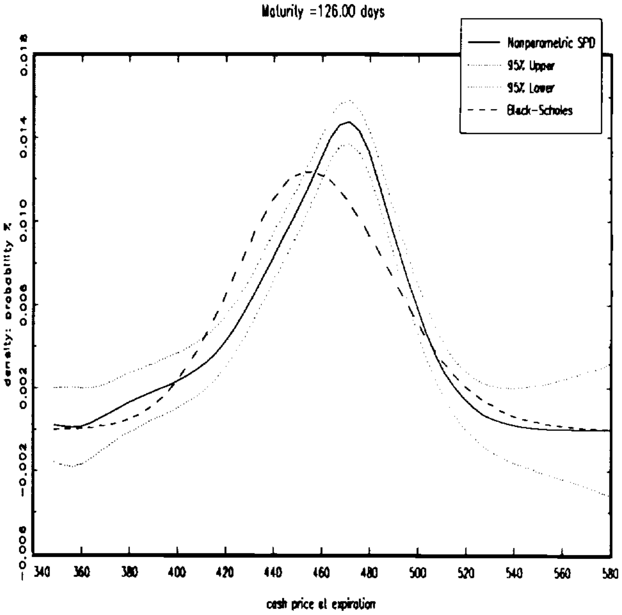


Figure 6
SPX Estimated Log-Return Densities
Maturity = 21.00 days

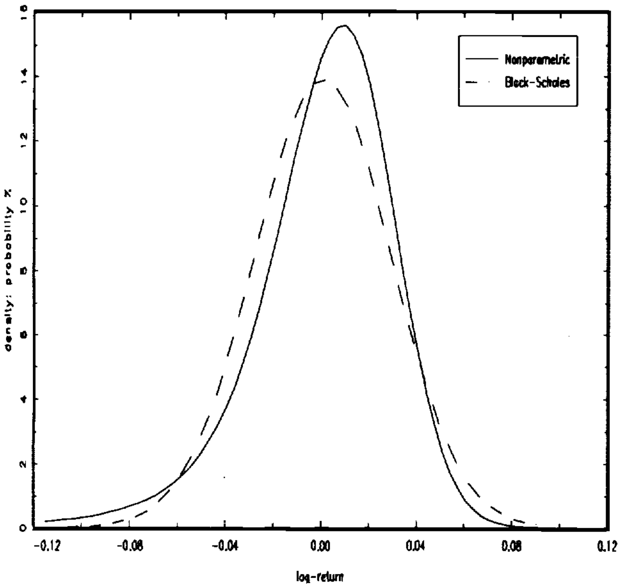


Figure 6
SPX Estimated Log-Return Densities
Maturity = 42.00 days

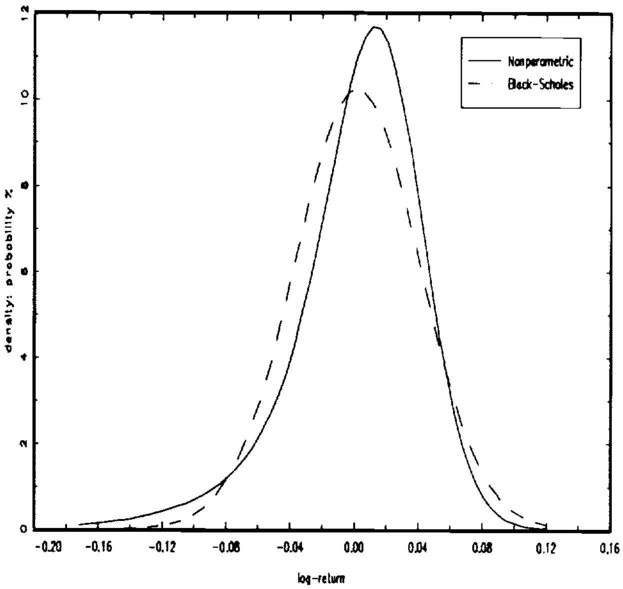


Figure 6
SPX Estimated Log-Return Densities
Maturity = 84.00 days

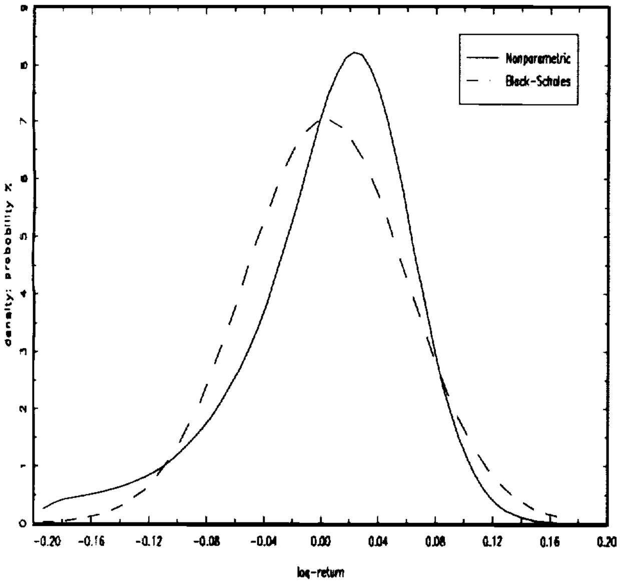


Figure 6
SPX Estimated Log-Return Densities
Maturity = 126.00 days

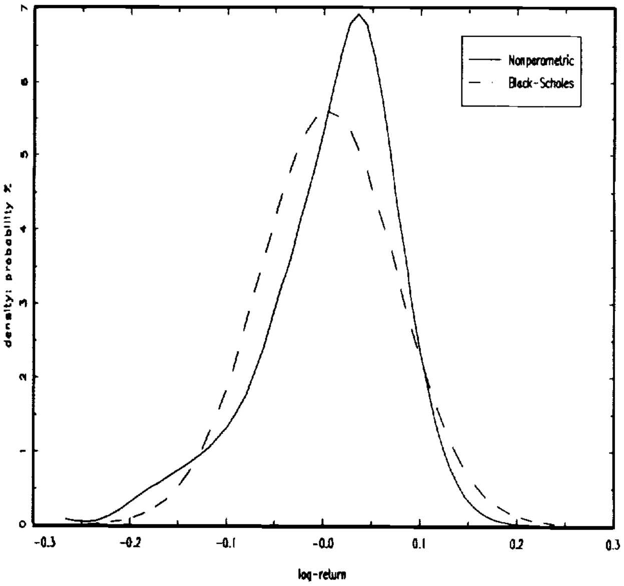


Figure 7a
Term Structure of Implied Mean of Log-Returns
(annualized)

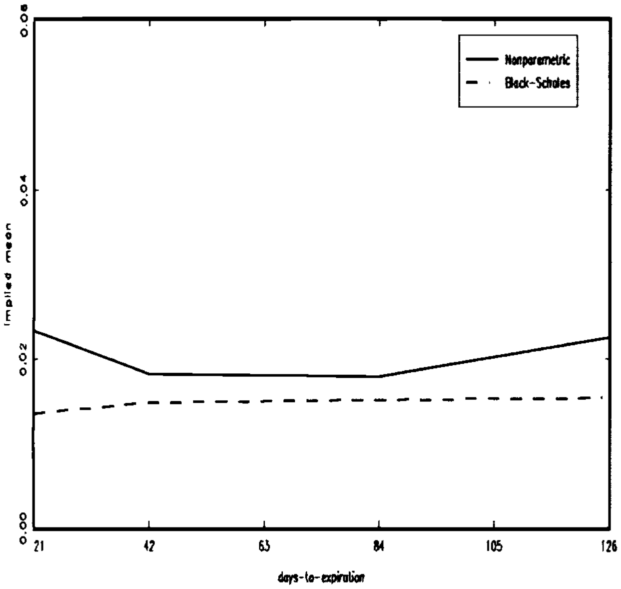


Figure 7b
Term Structure of Implied Volatility of Log-Returns
(annualized)

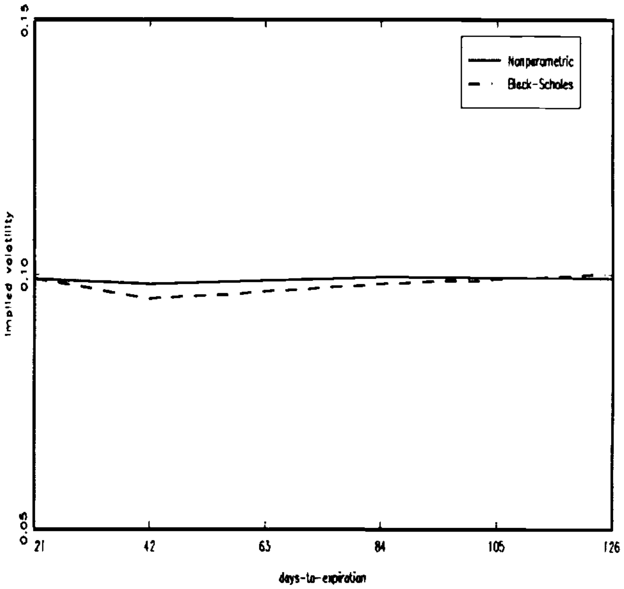
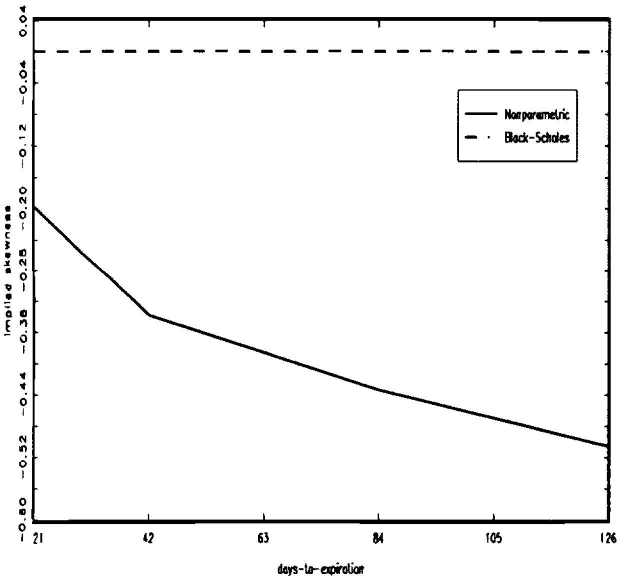
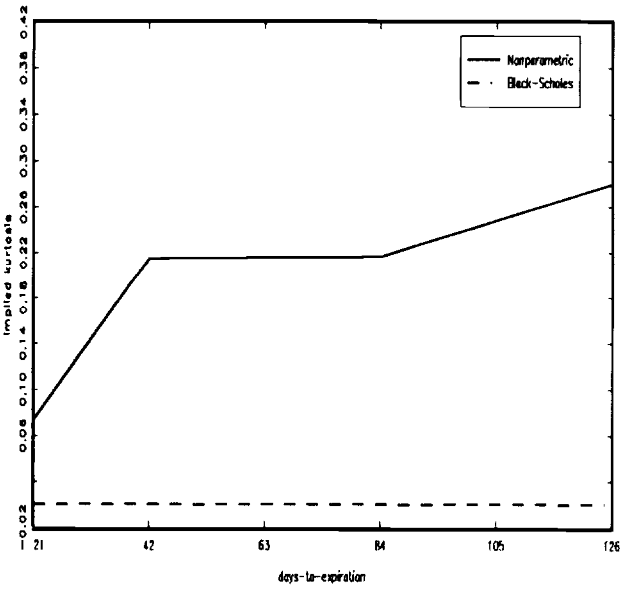


Figure 7c
Term Structure of Implied Skewness of Log-Returns
(annualized)



Figure 7d
Term Structure of Implied Kurtosis of Log-Returns
(annualized)