

Nonparametric Statistical Methods and the Pricing of Derivative Securities

RÜDIGER KIESEL[†]

r.t.kiesel@lse.ac.uk

Department of Statistics, London School of Economics and Political Science, Houghton Street, London WC2A 2AE, England

Abstract. In this review paper we summarise several nonparametric methods recently applied to the pricing of financial options. After a short introduction to martingale-based option pricing theory, we focus on two possible fields of application for nonparametric methods: the estimation of risk-neutral probabilities and the estimation of the dynamics of the underlying instruments in order to construct an internally consistent model.

Keywords: nonparametric methods, option pricing, exotic derivatives, term-structure of interest rates.

1. Introduction

The pricing and hedging of derivative securities is one of the key areas of stochastic finance and has attracted extensive attention and efforts from academic researchers (and clearly from financial practitioners as well) over the last decades. The key insights were gained in the early seventies by the Nobel-prize laureates Myron Scholes and Robert Merton, as well as the late Fischer Black, in their celebrated papers [8] and [34] (see also [35]) on pricing and hedging European call options (nowadays commonly called plain vanilla option, see §2 for formal definitions). However, ever since the introduction of the Black-Scholes-Merton (BSM) framework various shortcomings of their approach have been pointed out (see e.g. Fisher Black's account in 'The Holes in Black-Scholes', [7]) and various generalisations have been proposed, a recent overview of which is given in [44].

The advent of exotic options (sometimes called second- and third generation instruments, in contrast to the plain vanilla first generation options) has introduced a whole new dimension to pricing and hedging theory. While the BSM theory models certain underlying assets and produces prices of derivative securities based on these underlyings (nested in the

[†] Requests for reprints should be sent to Rüdiger Kiesel, Department of Statistics, London School of Economics and Political Science, Houghton Street, London WC2A 2AE, England

given stochastic model), the impact of the original assumptions on the underlying becomes more and more crucial as the complexity of the derivative asset increases. Indeed, the whole concept of an underlying has to be transformed. Traders will often price and hedge complex instruments using not the ‘old’ underlying but portfolios of the first generation options. Partly motivated by this change in practical attitude, we will focus on the so-called empirical option pricing approach, which uses information on plain vanilla options in order to generate a consistent pricing theory for exotic options.

Our aim is to supply the non-specialist reader with the necessary background knowledge in stochastic finance (texts on an introductory level are [5], [6]) and then to review and outline some of the statistical techniques recently employed to tackle questions arising from derivative security pricing. In particular we will concentrate on nonparametric estimation techniques to retrieve the relevant structure of the underlying without imposing restrictive parametric assumptions. We hope that this review will serve as an introduction to the range of existing problems arising from questions of valuation of financial derivatives and will motivate readers to explore the subject further (a systematic coverage of nonparametric and semiparametric methods in econometrics applied to various economically motivated questions is given in [37], while [12] is a comprehensive treatment of financial econometrics).

The paper is organised as follows: in §2 we give a brief introduction to the relevant concepts of stochastic finance, which we shall use later on. In particular we discuss the no-arbitrage pricing technique, the key concept in derivative pricing theory, the BSM-model for pricing options on stocks and the modelling of fixed-income securities. In §3 we outline the use of nonparametric methods for estimating risk-neutral probabilities from plain vanilla option prices and how these probabilities can be used to price exotic options, such as binary options. §4 is devoted to nonparametric estimation of drift and diffusion coefficients in diffusion models of the short rate. We conclude with a brief discussion of further developments.

2. Background on Modelling Financial Markets

2.1. Arbitrage Pricing and Equivalent Martingale Measures

The general underlying idea that we will follow is a relative pricing approach, i.e. we try to explain the prices of some assets in terms of other given and observable assets. This approach is based on the concept of arbitrage (or rather no arbitrage). Arbitrage pricing is independent of beliefs

and tastes (preferences) of the actors in the financial markets, with the only assumption being that all participants in the market simply prefer more to less. Therefore all investors constantly check whether they can restructure their current portfolios in such a way that they have to pay less for their restructured portfolios and still have the same (or even higher) return in the future. A situation like that is an arbitrage possibility – simply restructure the portfolio and pocket a riskless profit now – and should not be possible in an efficient market.

We model our financial market on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})$ (the filtration $\mathbb{F} = \{\mathcal{F}_t, 0 \leq t \leq T\}$, $\mathcal{F}_T = \mathcal{F}$ is used to model the flow of information in the market model) and assume a (finite) time horizon T . Taking the price processes, modelled as stochastic processes, of certain basic assets as given we can consider contingent claims modelled as \mathcal{F} -random variables (so we look at European claims with payoff at time T). We try to find the price processes for these claims based on the no-arbitrage principle. The link between no-arbitrage pricing and martingale theory has been developed in two fundamental papers by Harrison and Kreps in 1979 [23] and Harrison and Pliska in 1981 [24]. For standard stochastic models of financial markets (see [5] for details) the fundamental Harrison-Kreps-Pliska theorems are:

THEOREM 1 (FIRST FUNDAMENTAL ASSET PRICING THEOREM) *Call a probability measure \mathbb{P}^* equivalent to the actual probability measure \mathbb{P} an equivalent martingale measure if, under \mathbb{P}^* , the discounted (with respect to a suitable numéraire) price processes are (local) martingales. Then absence of arbitrage for the stochastic market model is equivalent to existence of an equivalent martingale measure \mathbb{P}^* .*

THEOREM 2 (SECOND FUNDAMENTAL ASSET PRICING THEOREM) *Call a market complete if every contingent claim can be replicated by a suitable portfolio of basic assets. Assuming that the model is arbitrage-free, then completeness is equivalent to uniqueness of the equivalent martingale measure.*

THEOREM 3 (RISK-NEUTRAL VALUATION) *Let X be a contingent claim in an arbitrage-free, complete market model. Then its price process $\Pi_X(t)$ is given by*

$$\Pi_X(t) = B(t)\mathbb{E}^* \left(\frac{X}{B(T)} \middle| \mathcal{F}_t \right), \quad (1)$$

where \mathbb{P}^* is the unique equivalent martingale measure, \mathbb{E}^* the expectation operator with respect to \mathbb{P}^* and B is the discount factor (numéraire).

In incomplete market situations we have several (infinitely many) equivalent martingale measures. For attainable contingent claims, i.e. contingent claims which can be replicated by a suitable portfolio of basic assets, the risk-neutral valuation formula (1) gives the same value whichever of the equivalent martingale measures is used. The question is which measure one has to choose to price the non-attainable claims. Usually some kind of utility-maximisation or ‘risk’-minimisation argument is given to motivate a particular choice (see [5], Chapter 7 for an introduction).

The above theorems are central to stochastic finance and have triggered various extensions and ramifications. Let us just mention a few sources for further study: [5], [16], [31], [36].

As we see from the above theorems the risk-neutral martingale measure plays the central role for pricing contingent claims. We will now consider two specific financial market models to explore it in more detail.

2.2. The Black-Scholes-Merton Model

We consider a financial market in which 2 assets are traded continuously up to a finite time horizon T . The first of these is called the bond, with price process defined by

$$dB(t) = B(t)r(t)dt, \quad B(0) = 1. \quad (2)$$

The remaining asset (usually referred to as stock) is subject to systematic risk. The price process $S(t), 0 \leq t \leq T$ of the stock is modelled by the linear stochastic differential equation

$$dS(t) = S(t)(b(t)dt + \sigma(t)dW(t)), \quad S(0) = p \in (0, \infty), \quad (3)$$

where $W(t), 0 \leq t \leq T$ is a standard one-dimensional Brownian motion. We assume that the underlying filtration \mathbb{F} is the Brownian filtration (basically $\mathcal{F}_t = \sigma(W(s), 0 \leq s \leq t)$ slightly enlarged to satisfy the usual conditions, compare [5], §5.8).

We interpret the process $r(t)$ as the interest-rate process - the short rate or the spot rate - i.e. $r(t)$ is the (instantaneously riskless) instantaneous interest rate. The process $b(t)$ is the appreciation rate for the stock, measuring the instantaneous rate of change of S at time t . Finally the volatility process $\sigma(t)$ models the instantaneous intensity with which the source of uncertainty influences the price of the stock at time t . The (stochastic) processes $r(t), b(t), \sigma(t)$ are referred to as the coefficients of the model. We assume that these coefficients satisfy certain measurability and integrability conditions (see e.g. [5], §6.2).

Our task is to price contingent claims on the stock within our model framework. Let us as before consider the example of a European call option on the stock S with strike K and expiry T . We can write the time T value of the option as

$$\max\{S(T) - K, 0\} = (S(T) - K)^+,$$

which is now a \mathcal{F} -measurable random variable, in fact a functional of $S(T)$. Black and Scholes in [8] and Merton [34] (see also [35], Chapter 9) constructed replicating portfolios to deduce the price of the call using arbitrage arguments.

Using the Harrison-Kreps-Pliska theory, i.e. Theorems 1 and 2, we see that (under regularity conditions on the coefficients) we have an arbitrage-free and complete model of a financial market (details can be found in [5] §6.2). Thus, by the risk-neutral valuation Theorem 3, the price process of any contingent claim X is given by

Proposition 1 *In a standard BSM-model the arbitrage price process of a contingent claim X is given by the risk-neutral valuation formula*

$$\Pi_X(t) = B(t)\mathbb{E}^* \left[\frac{X}{B(T)} \middle| \mathcal{F}_t \right] = \mathbb{E}^* \left[X e^{\{-\int_t^T r(u)du\}} \middle| \mathcal{F}_t \right]. \quad (4)$$

In the special case of the classical Black-Scholes model with constant coefficients r, b, σ we recover the famous Black-Scholes formula of a European call with strike K and maturity T on the stock S ,

Proposition 2 (Black-Scholes Formula) *The Black-Scholes price process of a European call is given by*

$$C(t) = S(t)N(d_1(S(t), T-t)) - Ke^{-r(T-t)}N(d_2(S(t), T-t)). \quad (5)$$

The functions $d_1(s, t)$ and $d_2(s, t)$ are given by

$$d_1(s, t) = \frac{\log(s/K) + (r + \frac{\sigma^2}{2})t}{\sigma\sqrt{t}},$$

$$d_2(s, t) = d_1(s, t) - \sigma\sqrt{t} = \frac{\log(s/K) + (r - \frac{\sigma^2}{2})t}{\sigma\sqrt{t}}$$

The statistical problems in using this formula are the estimation of the parameters, in particular the volatility σ , from historical time series of S . However, the particular assumptions on the dynamics of the underlying

process S have been questioned by several empirical investigations – see [12], §9.3 for an overview, [18] for an interesting recent study – and led to various extensions of the model, see e.g. stochastic volatility models, Lévy-type models, jump-diffusion models [20], [27], [35].

We will discuss in §3 nonparametric approaches to deducing information about the underlying process and indeed the risk-neutral martingale measure from market prices of options via the risk-neutral valuation formula (4).

2.3. Diffusion Models of the Term Structure of Interest rates

Let us now apply the ideas of §2.1 to the fixed-income securities market. As we want to develop a relative pricing theory, based on the no-arbitrage assumption, we will assume prices of some underlying objects as given. In the present context we take zero-coupon bonds as the building blocks of our theory. In doing so we face the additional modelling restriction that the value of a zero-coupon bond at time of maturity is predetermined ($= 1$). Furthermore, since the entirety of fixed-income securities gives rise to the term-structure of interest rates (sometimes called the yield curve), which describes the relationship between the yield-to-maturity and the maturity of a given fixed-income security, we face the further task of calibrating our model to a whole continuum of initial values (and not just to a vector of prices). A first attempt to explain the behaviour of the yield curve is in terms of a discrete set of spot rates of maturities between τ and T , where τ is the shortest (instantaneous) lending/borrowing period, and T the longest maturity of interest. We model these rates as correlated stochastic variables with the degree of correlation decreasing in terms of the difference in maturity. Empirical investigations show a very high degree of correlation of these rates. In fact using principal component analysis one can show that the first principal component often explains 80–90% of the total variance, and that the first three components taken together describe up to 90–95% of the total variance. Therefore, especially in early approaches, it has been tempting to choose one specific rate, usually the short rate, $r(t)$, as a proxy for the single variable that principal component analysis indicates can best describe the movements of the yield curve. The standard model is then of the form

$$dr(t) = a(r(t))dt + b(r(t))dW(t) \quad (6)$$

with functions a, b sufficiently regular (for instance, satisfying the conditions in [5] §5.7) and W a real-valued Brownian motion. Standard examples

are e.g. the *Vasicek* model $dr = (\alpha - \beta r)dt + \gamma dW$ and the *Cox-Ingersoll-Ross (CIR)* model $dr = (\alpha - \beta r)dt + \delta\sqrt{r}dW$.

Following the usual approach we fix an equivalent martingale measure \mathbb{P}^* (which we assume to exist), and model the short rate dynamics directly under \mathbb{P}^* . Thus we assume that r has \mathbb{P}^* -dynamics (6) with W a (real-valued) \mathbb{P}^* -Wiener process.

We can immediately apply the risk-neutral valuation technique to obtain the price process $\Pi_X(t)$ of any sufficiently integrable T -contingent claim X by computing the \mathbb{P}^* -expectation, i.e.

$$\Pi_X(t) = \mathbb{E}^* \left[e^{-\int_t^T r(u)du} X | \mathcal{F}_t \right], \quad (7)$$

in particular the price of a zero-coupon bond with maturity T should be given by

$$p(t, T) = \mathbb{E}^* \left[e^{-\int_t^T r(u)du} \mathbf{1} | \mathcal{F}_t \right], \quad (8)$$

Due to the incompleteness of the market implied by only specifying the short rate we now need to use (current) price information from the market in order to specify \mathbb{P}^* . This is done by *calibrating the model to data* or *fitting the yield curve*.

For practical purposes we would of course like to have an exact fit over the whole maturity spectrum. This involves an infinite-dimensional system of equations. So if we work with a model containing only a finite-dimensional parameter vector (such as the Vasicek model) there is no hope of obtaining a complete fit between observed and theoretical prices. This means that we are not able even to price the simplest contingent claims in this model (those with final payoff 1), let alone more complicated derivative structures. So for practical purposes one needs to introduce parameter functions of infinite dimensions to make calibration possible. For instance Hull and White [26] propose the *extended Vasicek* model $dr = (\alpha(t) - \beta(t)r)dt + \gamma(t)dW$ and the *extended CIR* model $dr = (\alpha(t) - \beta(t)r)dt + \delta(t)\sqrt{r}dW$. (with deterministic functions). Another natural approach, which stays in the above diffusion setting, is the nonparametric estimation of the coefficient functions a, b in (6), and we will discuss this approach in §4.

For further development and discussion of interest rate models we refer the reader to [5], [6], [36].

3. Options on Stocks

We now return to the setting in §2.2, i.e. we assume that a riskfree bond (now with constant interest rate $r(t) \equiv r$) and a risky stock are available for

trading. Assuming the existence of an equivalent martingale measure \mathbb{P}^* we know that the date- t prices of derivative securities with a single liquidation date- T payoff $X = \Phi(S(T))$ are given by the risk-neutral valuation formula (1), that is

$$\begin{aligned} \Pi_X(t) &= \mathbb{E}^* \left(e^{-\int_t^T r(u)du} \Phi(S(T)) \middle| \mathcal{F}_t \right) \\ &= e^{-r(T-t)} \int_0^\infty \Phi(x) f_t(x) dx, \end{aligned} \quad (9)$$

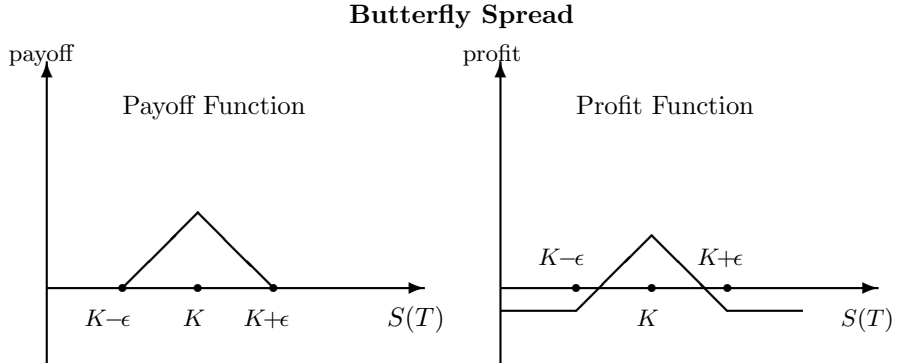
where $f_t(x)$ is the date- t density of the equivalent martingale measure (often called state-price density, in short: SPD) and r is the constant riskfree rate of interest between t and T . (Readers familiar with the subject will recognize that we performed a change of measure to the so-called T -forward measure.) In practice this corresponds to considering a European call option with a futures contract written on the asset as underlying. At maturity of the futures, the futures price equals the asset's spot price. Thus a European call option on the asset has the same value as a European call option on the futures contract with the same maturity. The effect of switching to the futures contract is that we are able to allow stochastic interest rates in our pricing approach. We refer the reader for further discussion of the T -forward measure and its applications to [5], §6.1 and [36] §13.2.2.

From a pricing perspective SPDs are 'sufficient statistics' in the sense that they summarize all relevant information about preferences and business conditions for the purposes of pricing financial derivatives. So it is very natural to try to infer SPDs from known and very liquid derivative prices and use them to price more sophisticated derivatives. Indeed, returning to our European call example Breeden and Litzenberger [10] have shown that it is possible to retrieve SPDs from call prices (see also Huang and Litzenberger [25] for a thorough discussion) and obtain the following formula:

$$e^{r\tau} \left. \frac{\partial^2 C}{\partial K^2} \right|_{K=x} = f_t^*(x). \quad (10)$$

The intuitive argument for obtaining formula (10) is as follows: Consider the portfolio obtained by selling two call options struck at K and buying one call struck at $K - \epsilon$ and one at $K + \epsilon$. Because of the shape of its payoff function $\Phi(S(T))$ (see figure) this portfolio is often called a butterfly spread:

(The fact that we have to invest to set up the portfolio is due to the convexity of call option prices in the strike, a property which follows from simple arbitrage considerations, see [5] §1.3 for details.) Consider now $1/\epsilon^2$ shares of the butterfly portfolio. Letting $\epsilon \rightarrow 0$ the payoff function of the butterfly tends to the Dirac delta function with unit mass at K , i.e. the



payoff of the butterfly is $1\mathcal{L}$ if $S(T) = K$ and nothing otherwise. Such a security is called an elementary Arrow-Debreu security and describes the relevant SPD (see [25]). Therefore the limit of $1/\epsilon^2$ shares of the butterfly portfolio should as ϵ tends to zero be equal to $e^{r\tau} f_t(X)$. So if we denote by $C(S(t), K, \tau)$ the market price of a call option at t with strike price X , time to maturity $\tau = T - t$ and underlying asset price $S(t)$, then

$$\frac{[2C(S(t), K, \tau) - C(S(t), K - \epsilon, \tau) - C(S(t), K + \epsilon, \tau)]}{\epsilon^2} \rightarrow \frac{\partial^2 C(S(t), K, \tau)}{\partial K^2}.$$

Of course we can verify the above by just formally differentiating the risk-neutral valuation formula for the European call. Indeed, in this case (1) is just

$$\begin{aligned} C(S(t), K, \tau) &= e^{-r\tau} \int_0^\infty \max\{x - K, 0\} f_t(x) dx \\ &= e^{-r\tau} \int_K^\infty (x - K) f_t(x) dx \end{aligned} \quad (11)$$

and differentiating with respect to K leads to (10).

Given (10) it is a straightforward idea to use European call option prices to estimate SPDs. Considering the well-documented shortcomings of parametric specification of stochastic asset price movements (which lead to parametric option pricing formulas) and the vast amount of available data one is quite naturally tempted to use nonparametric methods to estimate SPDs. The approach can be outlined in the following way (compare Aït-Sahalia and Lo [3]): use market prices to estimate the option pricing formula $\hat{C}(\cdot)$ nonparametrically, then differentiate this estimator twice with

respect to the strike price K to obtain $\partial^2 \hat{C} / \partial K^2$. Under suitable regularity assumptions the convergence in probability of \hat{C} to the true option pricing formula C implies that $\partial^2 \hat{C} / \partial K^2$ will converge to $\partial^2 C / \partial K^2$ which is proportional to the SPD. So we face the problem that given a set of historical option prices $\{C_i\}$ and accompanying characteristics $\{Z_i\} = (S(t_i), K_i, \tau_i, r_i)'$ (r_i is the constant interest rate from t_i to T) we seek a function $C(\cdot)$ that comes as close to C_i as possible. As a method of measuring closeness we use the quadratic distance, i.e. we solve

$$\min_{C \in \Gamma} \sum_{i=1}^n [C_i - C(Z_i)]^2$$

with Γ the class of all twice differentiable functions.

It is well-known that the best estimator is given by the conditional expectation of C with respect to Z . This leads to the use of nonparametric kernel regression methods.

The general problem of nonparametric estimation is to estimate relations such as

$$C_i = C(Z_i) + \epsilon_i, \quad i = 1, \dots, n \quad (12)$$

where $C(\cdot)$ is an unknown but fixed nonlinear function and $\{\epsilon_i\}_{i=1}^n$ is white noise ($\mathbb{E}(\epsilon|Z) = 0$, ϵ_i independent). The idea of kernel smoothers is to produce an estimate of C at $Z = z$ by giving more weight to observations (C_i, Z_i) with Z_i 'close' to Z . Therefore we must define what we mean by 'close' and 'far'. If we give too much weight to values in too large a neighborhood of z we will not be able to discover the genuine nonlinearities of $C(\cdot)$. On the other hand if we choose too small a neighborhood our estimate will not be able to differentiate noise from nonlinearities. To strike a careful balance between these two considerations we use a kernel function K , which acts as a weighting scheme (typically a probability density function, see [47], p. 12) and a smoothing parameter h which defines the degree of 'closeness' or neighborhood. The most widely used kernel estimator of a function $C(\cdot)$ as in (12) is the Nadaraya-Watson kernel estimator

$$\hat{C}_h(Z) = \mathbb{E}(C|Z) = \frac{\sum_{i=1}^n k\left(\frac{Z - Z_i}{h}\right) C_i}{\sum_{i=1}^n k\left(\frac{Z - Z_i}{h}\right)}. \quad (13)$$

The main issue now is to choose the kernel function $k(\cdot)$ and the bandwidth h in order to obtain the best possible fit. Although both $k(\cdot)$ and h have

an effect on the fit it is generally found in the literature that the choice of the smoothing parameter h is more important (see [22], [42], [47]). For the European option setting Aït-Sahalia and Lo [3] determine the bandwidth h using cross-validation techniques (see [22], chap. 3 for a discussion of this technique) to optimally balance the kernel estimator and to address the practical problems discussed below (see also [21] for a discussion and more references on the choice of the smoothing parameter and [11] for a related approach to American option pricing).

In the option pricing setting the kernel approach allows the following intuition: The price of an European call option with characteristics Z is given as a weighted average of observed prices C_i with more weight given to the options whose characteristics Z_i s are closer to the characteristics Z of the option to be priced. For practical applications involving option price data we can't maintain the independence assumption on the error term, since we have to deal with nonstationary correlated data. However, it is found that the variable S/K is stationary as the exercise prices bracket the underlying price process. The issue of correlation of the data is more involved. Under general conditions the kernel estimator remains convergent and asymptotically normal and only the asymptotic variance is affected by the correlation of the error terms (see [1]), but there is no general result on how to avoid over- or undersmoothing (see [15] for an overview and [?] for some simulation results on special methods). In our context the complexity of the problem is even increased by the fact that at each time t several option contracts are listed simultaneously and trading may take place in only a subset of these contracts. Additionally we face the so-called 'curse of dimensionality' problem, i.e. the precision of the estimates deteriorate as the dimension of Z increases, unless the sample size increases drastically (see [42] chap. 7). This is particular important for our approach since we need to differentiate the call pricing function twice with respect to K to obtain the state price density. (Observe that we also can find estimators for the Delta of the option, needed for hedging purposes, by differentiating the call price function with respect to S). To reduce the number of regressors we could

- (i) impose regularity condition on the option pricing formula to assume that it is homogeneous of degree one in S and K (as is the Black-Scholes formula);
- (ii) assume that the option pricing formula is not a function of the asset price S and the riskfree rate r separately, but only depends on these variables through the futures price $F_\tau = Se^{r\tau}$ (by the martingale property the mean of the SPD is F_τ and the assumption would be the whole

distribution depends only on the futures price – as is the case in the BMS-model);

- (iii) use a semiparametric-approach, i.e. suppose that the option pricing formula is given by a certain parametric structure except that some parameters are given as nonparametric functions, which we need to estimate.

We will only consider the semiparametric approach using the Black-Scholes formula in more detail. We suppose that the call price function is given by the parametric Black-Scholes formula except that the implied volatility parameter for that option is a nonparametric function $\sigma(S(t), K, \tau)$, so:

$$C(S(t), K, \tau, r) = C_{BS}(S(t), K, \tau; \sigma(S(t), K, \tau)) \quad (14)$$

In this semiparametric model we only need to compute the three-dimensional regression of implied volatilities on asset price, strike and time-to-expiration to estimate $\mathbb{E}(\sigma|S(t), K, \tau)$.

We now describe the full-dimensional and the Black-Scholes induced semiparametric model. Given the data $\{C_i, S(t_i), K_i, r_i, \tau_i\}$ the full-dimensional nonparametric SPD estimator is formed using a multivariate kernel in (13) formed as a product of $d = 4$ univariate kernels

$$\begin{aligned} & \hat{C}(S(t), K, \tau, r_\tau) \\ &= \frac{\sum_{i=1}^n k_S \left(\frac{S(t) - S(t_i)}{h_S} \right) k_K \left(\frac{K - K_i}{h_K} \right) k_\tau \left(\frac{\tau - \tau_i}{h_\tau} \right) k_r \left(\frac{r_\tau - r_i}{h_r} \right) C_i}{\sum_{i=1}^n k_S \left(\frac{S(t) - S(t_i)}{h_S} \right) k_K \left(\frac{K - K_i}{h_K} \right) k_\tau \left(\frac{\tau - \tau_i}{h_\tau} \right) k_r \left(\frac{r_\tau - r_i}{h_r} \right)} \end{aligned}$$

In the semiparametric model we have to form the three-dimensional kernel estimator of $\mathbb{E}(\sigma|S(t), K, \tau)$ as

$$\hat{\sigma}(S(t), K, \tau) = \frac{\sum_{i=1}^n k_S \left(\frac{S(t) - S(t_i)}{h_S} \right) k_K \left(\frac{K - K_i}{h_K} \right) k_\tau \left(\frac{\tau - \tau_i}{h_\tau} \right) \sigma_i}{\sum_{i=1}^n k_S \left(\frac{S(t) - S(t_i)}{h_S} \right) k_K \left(\frac{K - K_i}{h_K} \right) k_\tau \left(\frac{\tau - \tau_i}{h_\tau} \right)}$$

where σ_i is the implied volatility corresponding to the option price C_i . Then the call pricing function is estimated as

$$\hat{C}(S(t), K, \tau, r_\tau) = C_{BS}(S(t), K, \tau, r_\tau; \hat{\sigma}(S(t), K, \tau)). \quad (15)$$

In both cases the univariate kernel functions k , and the bandwidth parameters h , are chosen to optimize the asymptotic properties of the second derivative of \hat{C} . We obtain then the SPD estimator by taking the second partial derivative of $\hat{C}(\cdot)$ with respect to K and setting

$$f_t(x) = e^{r\tau} \left[\frac{\partial^2 \hat{C}(S(t), K, \tau, r\tau)}{\partial K^2} \right]_{K=x}.$$

Let us mention again that we also obtain an estimator for the option's delta via

$$\Delta(t) = \frac{\partial \hat{C}(S(t), K, \tau, r\tau)}{\partial S(t)}. \quad (16)$$

Aït-Sahalia and Lo [3] conduct a Monte-Carlo study of the performance of their nonparametric approach to estimate the call option pricing function and the state-price density. They simulate 5000 replication of a trading year (with on average 80 call options per trading day) based on a Black-Scholes model (i.e. geometric Brownian motion for the underlying and option prices computed according to the BMS-formula). Their findings are that their nonparametrically estimated quantities are within one percent of their theoretical counterparts, and thus virtually free of any bias.

Any practical implications of option pricing models must be concerned with matters such as illiquidity, dividend corrections, bid-ask spreads, timing issues and other market microstructure issues.

The dataset used in [3] are S&P500 index option prices. The S&P500 index option is among the most actively traded financial derivatives in the world and using an index ensures the highest likelihood that the Black-Scholes assumptions are actually satisfied. Furthermore, since all option prices are recorded at the same time on each day, only the index price has to temporally matched. To correct for dividends they infer futures prices for each maturity (which removes the unobservable dividend rate) by put-call parity. However, since in-the-money derivatives are very illiquid they only use at-the-money calls/puts (that is they only have a limited range of strike prices available). Finally, given the futures prices all illiquid call prices can be replaced via put-call parity (for illiquid calls puts must be out-of-the money and hence liquid).

Let us now mention a few related approaches:

Jackwerth and Rubinstein in a series of papers [29], [30], [41] construct implied binomial trees, in which risk-neutral probabilities $\{p_i^*\}$ associated with the binomial terminal stock price $S(T)$ are estimated by minimizing the sum of squared deviations between $\{p_i^*\}$ and a set of prior risk-neutral probabilities $\{\tilde{p}_i^*\}$ subject to the restriction that $\{p_i^*\}$ correctly price an existing set of options and the underlying stock. Since any state price density

(SPD) can be approximated to any degree of accuracy by a binomial tree this approach can be viewed as nonparametric. The main difference from kernel estimation is that one set of risk-neutral probabilities $\{p_i^*\}$ is fitted for each cross-section of options (so a n -step implied binomial tree at t_1 will be different from a n -step tree at t_2), whereas kernel SPDs aggregate option data over time to get a single SPD (n -step SPDs are the same for t_1 and t_2). Therefore while implied binomial trees are completely consistent with all option prices at each date, they are not necessarily consistent over time. In contrast the nonparametric kernel SPD is consistent across time, but there may be dates for which the SPD fits the cross section of option data poorly. Which method to prefer is therefore a question of the modelling strategy and purpose. Methods similar to Rubinstein's approach, i.e. fitting risk-neutral density functions using some general family of distributions, are discussed in [?], [19], [33], [40]. Rady [39] considers the underlying problem in a general framework: given a valuation formula $u(S(t), t)$ for a derivative asset, what can be inferred about the underlying asset price process? Such a result has applications to nonparametric approaches to derivative asset pricing via learning networks, see [28]. Once a pricing formula has been learnt by the network, the risk-neutral law of the underlying process can be identified.

Stutzer [46] develops an option pricing framework entirely based on the empirical distribution of the historical time series of the returns of the underlying and uses the maximum entropy principle to choose the risk-neutral probability measure needed for pricing purposes. Such an approach does not require the use of derivative security prices and thus the various complications implied by such data are avoided. Furthermore it might provide a more accurate assessment of the likely pricing impact caused by investors' data-based beliefs about the future value distribution.

One of the first semiparametric approaches has been proposed by Shimko [43]. He applies equation (14) with the (implied) volatility function only depending on the parameters 'strike' and time to maturity, that is he uses $\sigma(S(t), K, \tau) \equiv \sigma(K, \tau)$, where the volatility function is assumed to be a parabolic function $\sigma(K) = a_0(\tau) + a_1(\tau)K + a_2(\tau)K^2$ (the choice being motivated by the 'smile-effect' observable in the profile of implied volatility across strike values). After obtaining a least-squares fit for the volatility function it can be plugged into the Black-Scholes formula and the state-price density is obtained by straight-forward differentiation.

Finally, Broadie et al. [11] apply nonparametric techniques to the pricing of American-type options. Recall that for American-type options early exercise is possible, i.e. it is possible to exercise the option before the maturity date T . The optimal strategy leads to an early exercise boundary

which triggers exercise of the option. Broadie et al. develop nonparametric estimators for the early exercise boundary and option prices and compare them with their parametric analogues. They find that there exist significant discrepancies between the two methods.

We close with an application to the pricing of a particular type of exotic options, the so-called binary options (see [5], §6.3.5). A binary (or digital) option is a contract whose payoff depends in a discontinuous way on the terminal price of the underlying asset. The most popular variant is the *Cash-or-nothing* option. Here the payoffs at expiry of the European call resp. put are given by

$$BCC(T) = C\mathbf{1}_{\{S(T) > K\}} \quad \text{resp.} \quad BCP(T) = C\mathbf{1}_{\{S(T) < K\}}.$$

By the risk-neutral valuation principle (1) we can price these derivatives by calculating the conditional expectation over the final payoff. For the cash-or-nothing option we get

$$\begin{aligned} BCC(T) &= e^{-r\tau} \mathbb{E}^*(C\mathbf{1}_{\{S(T) > K\}} | \mathcal{F}_t) \\ &= e^{-r\tau} C \int_K^\infty f_t(x) dx = e^{-r\tau} C(1 - F_t(K)), \end{aligned}$$

where $F_t(\cdot)$ is just the cumulative distribution function related to the SPD. Ait-Sahalia and Lo [3] compare cash-or-nothing calls priced nonparametrically with the corresponding BSM prices. Not surprisingly they find higher out-of- and in-the-money and lower at-the-money nonparametrically estimated prices, a fact that can be explained by the ability of the nonparametric estimator to accommodate for skewness and kurtosis found in financial time series.

4. The Term-Structure of Interest Rates

We consider estimation of the relevant parameters under the assumption that the short rate is given as a time-homogeneous diffusion

$$dr(t) = a(r(t))dt + b(r(t))dW(t). \quad (6)$$

We want to estimate the functions a and b without specifying any parametric assumptions about either the drift or the diffusion. For the purpose of option pricing we furthermore need to specify an equivalent martingale measure. In practical applications this is usually done by specifying or estimating the so-called market price of risk λ , which quantifies the additional expected rate of return on a bond over the riskfree rate per additional unit

of risk (expressed in terms of the volatility) that an investor requires (see [5], §8.2). Since it is hard to form an a priori idea of the functional form of the diffusion function and long time series of daily data on spot interest rates are typically available, a nonparametric approach is appropriate. We will now discuss three specific nonparametric procedures, which have been proposed recently.

4.1. A Density-Matching Approach

In [1] Ait-Sahalia uses a density-matching approach to derive the instantaneous drift and diffusion functions of the short rate process so as to be consistent with the observed distribution of the discrete data. Since the process defined by the stochastic differential equation (6) is fully characterised by the first two moments, i.e. the drift and diffusion function, the form of these functions fully specify the marginal and transitional density of the spot rate process $r(t)$ and vice versa.

The idea of the drift matching approach is to estimate nonparametrically the transitional densities of the short rate process and then find drift and diffusion processes matching these densities. The mathematical motivation is as follows:

Use the Kolmogorov forward equation for the transition densities $p(s, y|t, x)$ from value x at time t to value y at time s

$$\frac{\partial p(s, y|t, x)}{\partial s} = -\frac{\partial}{\partial y}(a(y)p(s, y|t, x)) + \frac{1}{2} \frac{\partial^2}{\partial y^2}(b^2(y)p(s, y|t, x)). \quad (17)$$

Letting $t \rightarrow -\infty$ in (17) we obtain

$$2 \frac{d}{dx}(a(x)\pi(x)) = \frac{d^2}{dx^2}(b^2(x)\pi(x)), \quad (18)$$

with π the stationary density. Integration yields

$$a(x) = \frac{2}{\pi(x)} \frac{d}{dx}(b^2(x)\pi(x)) \quad (19)$$

and we can use (19) to find the remaining function given any two.

Ait-Sahalia [1] assumes that the drift coefficient $a(x)$ has the parametric form $a(x) = k(\theta - x)$ and estimates the drift by ordinary least squares. Observe that $\mathbb{E}(r(t + \Delta)|\mathcal{F}_t) = \theta(1 - e^{-k\Delta}) + e^{-k\Delta}r(t)$ independent of the diffusion coefficient σ (compare [5], §5.7 and Exercise 5.9 for a related computation). Then b can be estimated nonparametrically by integrating

(19) once more to obtain

$$b^2(x) = \frac{2}{\pi(x)} \int_0^x a(u)\pi(u)du. \quad (20)$$

Technically the procedure is as follows: Start with interest rate data $\{r_i, i = 1, \dots, n\}$ and form a smooth density estimator

$$\hat{\pi}(r) = \left(\frac{1}{nh_n} \sum_{i=1}^n K((r - r_i)/h_n) \right) \quad (21)$$

based on a kernel function $K(\cdot)$ and bandwidth h_n (compare §3 for the technical details of kernel smoothing). Estimate the parameters of the drift function $a(\cdot, \theta)$ by ordinary least squares (OLS); call the estimator \hat{a} . Now use the kernel estimator $\hat{\pi}$ for the transition densities π and the drift estimator \hat{a} in (20) to get an estimator \hat{b} for the diffusion function b . Use \hat{b} to correct for heteroskedasticity in the regression and construct the weighting matrix for the second step feasible generalized least squares (FGLS).

To compute prices of interest-rate-derivative securities the market price of risk λ has to be specified (in order to find the relevant risk-neutral martingale measure in (1)). Aït-Sahalia assumes the market price of risk is constant and estimates it by minimising the squared deviation between a given yield curve and that implied by the model (in effect minimising differences of bond prices).

4.2. Direct Nonparametric Estimation of the Diffusion Coefficient

Bianchi et al. [4] assume (like Aït-Sahalia [1]) a parametric model of the drift and try to estimate the diffusion function nonparametrically. In contrast to Aït-Sahalia they applied direct nonparametric regressions to the instantaneous variance rather than calculating the integral of a mean function with respect to a nonparametrically estimated density. Therefore the approach needs to discretise the model first in order to estimate the diffusion function, which leads to the implicit assumption that more data means more frequent data on a fixed period of observation. This condition is hardly matched in practice, so one has to be careful with arguments based on asymptotic properties of the estimator.

Technically the approach [4] proceeds as follows: assuming an affine structure of the drift term, i.e.

$$a(r) = \alpha + \beta r \quad (22)$$

(so the usual parametrisation in the mean-reverting approach is $a(r) = \kappa(r^* - r)$ with $r^* = -\alpha/\beta, \kappa = -\beta/\Delta t$), one can use a first-order discretisation (see e.g. [32] for a discussion of discretisation methods)

$$\Delta r_{t+\Delta t} = (\alpha + \beta r_t)\Delta t + u_{t+\Delta t}\sqrt{\Delta t}. \quad (23)$$

From the OLS estimates of α and β , $\hat{\alpha}$ and $\hat{\beta}$, one can derive $\hat{\kappa}$ and \hat{r}^* , then

$$\hat{u}_{t+\Delta t}\sqrt{\Delta t} = \Delta r_{t+\Delta t} - \hat{\kappa}(\hat{r}^* - r_t) \quad (24)$$

and a proxy for actual volatility at time $t + \Delta t$ is given by

$$\hat{\epsilon}_{t+\Delta t} = \frac{1}{\Delta t}\hat{u}_{t+\Delta t} \quad t = 1, \dots, T-1. \quad (25)$$

From the $\hat{\epsilon}$'s the diffusion function $b(r)$ can be estimated nonparametrically by smoothing the scatterplot of $\hat{\epsilon}_{t+\Delta t}^2$ against r_t , that is, using

$$\mathbb{E}(\hat{\epsilon}_j^2) = g(r_j), \quad (26)$$

where the $\hat{\epsilon}_j^2$ are the residuals $\hat{\epsilon}_{t+\Delta t}^2$ sorted according to the sorting (in ascending order) of r_t .

The unknown smooth function $g(r) = b^2(r)$ can be estimated by any nonparametric smoothers (see [22], [42], [47]) such as smoothing splines, kernel regressors or local polynomial estimators. After obtaining the drift and diffusion function the pricing of interest-rate derivatives is done by the PDE equivalent to risk-neutral valuation formula (1) (see [5], chapter 8), with the market price of risk $\lambda = 0$, i.e. it is assumed that the data are observed under an equivalent martingale measure.

4.3. A Fully Nonparametric Approach

Stanton [45] proposes a fully nonparametric approach by inverting a numerical approximation scheme. Considering a diffusion process $X(t)$, which satisfies the stochastic differential equation (6), that is $dX(t) = a(X(t))dt + b(X(t))dW(t)$, he proceeds as follows:

Step 1: Under suitable restriction on a, b and a further function f we have the following Taylor Series approximation (compare [32]):

$$\begin{aligned} \mathbb{E}_t(f(X_{t+\Delta}, t + \Delta)) &= f(X_t, t) + \mathcal{L}f(X_t, t)\Delta + \frac{1}{2}\mathcal{L}^2 f(X_t, t)\Delta^2 + \\ &+ \dots + \frac{1}{n!}\mathcal{L}^n f(X_t, t)\Delta^n + \mathbf{O}(\Delta^n), \end{aligned}$$

where \mathcal{L} is the infinitesimal generator of the diffusion $X(t)$.

Step 2: Now use suitable functions f in order to construct approximations of a and b .

For instance using $f(x, t) = x$ gives $\mathcal{L}f(x, t) = a(x)$ hence the above leads to an approximation for the drift, or using $g(x, t) = (x - X(t))^2$ we find $\mathcal{L}g(x, t) = 2(x - X(t))a(x) + b^2(x)$ and so $\mathcal{L}g(X(t), t) = b^2(X(t))$.

The first-order estimates following this approach are

$$a(X(t)) = \frac{1}{\Delta} \mathbb{E}(X(t + \Delta) - X(t) | \mathcal{F}_t) + \mathbf{O}(\Delta) \quad (27)$$

and

$$b^2(X(t)) = \frac{1}{\Delta} \mathbb{E}((X(t + \Delta) - X(t))^2 | \mathcal{F}_t) + \mathbf{O}(\Delta). \quad (28)$$

The market price of risk λ is estimated similarly using two risky assets (i.e. bonds with different maturities, compare §2.3) with the first-order approximation being

$$\lambda(r(t)) = \frac{b(r(t))}{\Delta(b^{(1)}(r(t)) - b^{(2)}(r(t)))} \mathbb{E}(R_{t,t+\Delta}^{(1)} - R_{t,t+\Delta}^{(2)} | \mathcal{F}_t) + \mathbf{O}(\Delta), \quad (29)$$

where $b^{(i)}$ and $R_{t,t+\Delta}^{(i)}$ are the instantaneous volatility and holding period return of asset $i = 1, 2$ respectively.

Finally, Stanton applies the Nadaraya-Watson kernel regression estimator (13) with a Gaussian kernel to estimate the conditional expectations in equations (27), (28) and (29).

4.4. Related Topics

Chan et al. [13] present an empirical discussion of alternative (parametric) models of the short-term interest rate, which might be extended to include the nonparametrically estimated models above. A first step in this direction was done by Pritsker [38]. There the Vasicek-model is used as an example for a persistent time series, which highlights some problematic aspects of nonparametric approaches to spot-rate modelling. In particular it is shown that persistence has an important impact on optimal bandwidth selection and finite sample performance (usual nonparametric approaches do not account for that). Results in [2] (rejection of most major spot-rate models) are shown to be very sensitive to these features. Further critical aspects of the finite sample performance of the Ait-Sahalia and Stanton results are in Chapman and Pearson [14]. Chapman and Pearson conduct a Monte Carlo study based on a Cox-Ingersoll-Ross model and find that both nonparametric estimators suggest a nonlinear drift term.

Downing [17] contains a further discussion of the nonparametric approach. He applies the nonparametric estimation techniques proposed in [9] to simulated data from a two-factor interest rate model. He found that the proposed nonparametric estimator exhibits a significant bias and a high sampling variance. Based on these result his conclusion is that nonparametric estimators are most useful as diagnostic tools rather than for actual estimation of drift and diffusion functions.

5. Conclusion

The rapidly increasing degree of sophistication of financial products (especially derivative products) has emphasised the need for accurate stochastic modelling of the price processes of various financial assets. Since parametrically restricted models have often been found violated by empirical studies there has been an intense interest in nonparametric modelling approaches. Although highly data-intensive, generally requiring several thousand data-points for a reasonable level of accuracy, the very structure of financial markets producing a massive amount of data invites the application of these methods. Furthermore, nonparametric methods are adaptive, responding to structural shifts in the data-generating process, and are thus more reliable when used for risk-management. Their flexibility allows for modelling a wide range of derivative securities and underlying asset dynamics, of which we presented the two most important examples. Finally the implementation of nonparametric methods is comparably simple and allows for computationally efficient implementation.

A similar discussion to ours with more emphasis on the statistical side is given in [21].

Acknowledgments

The author likes to thank N.H.Bingham and Uma Moorty for many helpful and encouraging comments.

References

1. Y. Ait-Sahalia. Nonparametric pricing of interest rate derivative securities. *Econometrica*, 64(3):527–600, 1996.
2. Y. Ait-Sahalia. Testing continuous time models of the spot interest rate. *Review of Financial Studies*, 2(9):385–426, 1996.
3. Y. Ait-Sahalia and A.W. Lo. Nonparametric estimation of state-price densities implicit in financial asset prices. *Journal of Finance*, 53(3), 1998.

4. M. Bianchi, M. Orszag, and J. Steeley. Semi-parametric modelling of the term structure. Paper presented at the CEPR/ESRC/IFR Finance Network Workshop: Financial Econometrics, London, March 1998, 1998.
5. N. Bingham and R. Kiesel. *Risk-neutral valuation*. Springer Finance. Springer London, 1998.
6. T. Björk. *Arbitrage theory in continuous time*. Oxford University Press, 1999.
7. F. Black. The holes in Black-Scholes. *RISK*, 1(4), March 1988. reprinted in: A Nobel model, *RISK*, Black-Scholes-Merton supplement, September 1998.
8. F. Black and M. Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 72:637–659, 1973.
9. J. Boudoukh, M. Richardson, R. Stanton, and R.F. Whitelaw. The stochastic behaviour of interest rates: Implications from a multifactor, nonlinear continuous-time model. Working Paper, Haas School of Business, 1999.
10. D. Breeden and R. Litzenberger. Prices of state-contingent claims implicit in option prices. *Journal of Business*, 51:621–615, 1978.
11. M. Broadie, J. Detemple, E. Ghysels, and O. Torr es. Nonparametric estimation of American options exercise boundaries and call prices. Discussion paper, CIRANO, Montr al, 1995.
12. J.Y. Campbell, A.W. Lo, and A.C. MacKinlay. *The econometrics of financial markets*. Princeton University Press, 1997.
13. K.C. Chan et al. An empirical comparison of alternative models of the short-term interest rates. *Journal of Finance*, 47:1209–1228, 1992.
14. D.A. Chapman and N.D. Pearson. Is the short rate drift actually nonlinear? Forthcoming in the *Journal of Finance*, 1999.
15. R. Chen, W. Hardle, and H. Lutkepohl. A review of nonparametric time series analysis. Humboldt Universitat zu Berlin, August 1996.
16. F. Delbaen and W. Schachermayer. A general version of the fundamental theorem of asset pricing. *Mathematische Annalen*, 300:463–520, 1994.
17. C. Downing. Nonparametric estimation of multifactor continuous time interest rate models. Working Paper, Federal Reserve Board, 1999.
18. B. Dumas, J. Fleming, and R.E. Whaley. Implied volatility functions: Empirical tests. *Journal of Finance*, 53(6), 1998.
19. B. Dupire. Pricing and hedging with smiles. In M.A.H. Dempster and S.R. Pliska, editors, *Mathematics of derivative securities*, pages 102–111. Cambridge University Press, 1997.
20. E. Eberlein and U. Keller. Hyperbolic distributions in finance. *Bernoulli*, 1:281–299, 1995.
21. E. Ghysels et al. Non-parametric methods and option pricing. In D.J. Hand and S.D. Jacka, editors, *Statistics in finance*, Arnold Applications of Statistics, chapter 13, pages 261–282. Arnold, London, 1998.
22. W. Hardle. *Applied nonparametric regression*, volume 19 of *Econometric Society Monographs*. Cambridge University Press, 1990.
23. J.M. Harrison and D. M. Kreps. Martingales and arbitrage in multiperiod securities markets. *J. Econ. Th.*, 20:381–408, 1979.
24. J.M. Harrison and S.R. Pliska. Martingales and stochastic integrals in the theory of continuous trading. *Stochastic Processes and their Applications*, 11:215–260, 1981.
25. C-F. Huang and R.H. Litzenberger. *Foundations for financial economics*. Prentice Hall, Englewood Cliffs, New Jersey, 1988.
26. A. Hull and A. White. Pricing interest rate derivative securities. *Review of Financial Studies*, 3(4):573–592, 1990.
27. J. Hull. *Options, futures, and other derivative securities*. Prentice-Hall, 3rd edition, 1997. 2nd ed. 1993, 1st ed. 1989.

28. J.M. Hutchinson, A.W. Lo, and T. Poggio. A nonparametric approach to pricing and hedging derivative securities via learning networks. *Journal of Finance*, 49(3):851–889, 1994.
29. J. Jackwerth and M. Rubinstein. Recovery probability distributions from option prices. *Journal of Finance*, 51:1611–1631, 1996.
30. J. Jackwerth and M. Rubinstein. Recovering probabilities and risk aversion from option prices and realized returns. Preprint, London Business School, <http://www.lbs.ac.uk/faculty/jjackwerth>, September 1997.
31. J. Jacod and A.N. Shiryaev. Local martingales and the fundamental asset pricing theorems in the discrete-time case. *Finance and Stochastics*, 2(3):259–273, 1998.
32. P.E. Kloeden and E. Platen. *Numerical solutions of stochastic differential equations*, volume 23 of *Applications of Mathematics, Stochastic Modelling and Applied Probability*. Springer Verlag, Berlin Heidelberg New York, 1992.
33. F.A. Longstaff. Option pricing and the martingale restriction. *Review of Financial Studies*, 8(4):1091–1124, 1995.
34. R.C. Merton. Theory of rational option pricing. *Bell Journal of Economics and Management Science*, 4:141–183, Spring 1973.
35. R.C. Merton. *Continuous-time finance*. Blackwell, Oxford, 1990.
36. M. Musiela and M. Rutkowski. *Martingale methods in financial modelling*, volume 36 of *Applications of Mathematics: Stochastic Modelling and Applied Probability*. Springer, New York, 1997.
37. A. Pagan and A. Ullah. *Nonparametric econometrics*. Themes in modern econometrics. Cambridge University Press, 1999.
38. M. Pritsker. Nonparametric density estimation and tests of continuous time interest rate models. FRB working paper, 1997.
39. S. Rady. State prices implicit in valuation formula for derivative securities. Discussion Paper No. 181, LSE Financial Market Group, October 1994.
40. J. Rosenberg. Pricing contingent claims using estimated risk-neutral density functions. New York University, Stern School of Business, Working Paper FIN-96-36, 1996.
41. M. Rubinstein. Implied binomial trees. *Journal of Finance*, 49(3):771–818, 1994.
42. D.W. Scott. *Multivariate density estimation: Theory, practice, and visualization*. John Wiley & Sons, 1992.
43. D. Shimko. Bounds of probability. *Risk*, 6, 1993.
44. C. Smithson and S. Song. Extended family. *RISK*, September 1998. In: A Nobel model, Black-Scholes-Merton supplement, p. 14–17.
45. R. Stanton. A nonparametric model of the term structure dynamics and the market price of interest rate risk. *Journal of Finance*, 52(5):1973–2002, 1997.
46. M. Stutzer. A simple nonparametric approach to derivative security valuation. *Journal of Finance*, 51(5):1633–1652, 1996.
47. M.P. Wand and M.C. Jones. *Kernel Smoothing*, volume 60 of *Monographs on Statistics and Applied Probability*. Chapman & Hall, 1995.