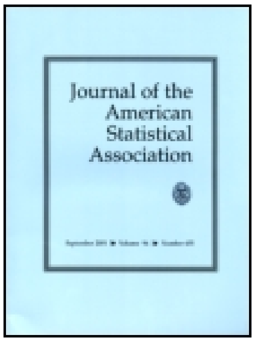

Journal of the American Statistical Association

ISSN: 0162-1459 (Print) 1537-274X (Online) Journal homepage: www.tandfonline.com/journals/uasa20

Option Pricing With Model-Guided Nonparametric Methods

Jianqing Fan & Loriano Mancini

To cite this article: Jianqing Fan & Loriano Mancini (2009) Option Pricing With Model-Guided Nonparametric Methods, Journal of the American Statistical Association, 104:488, 1351-1372, DOI: 10.1198/jasa.2009.ap08171

To link to this article: https://doi.org/10.1198/jasa.2009.ap08171

Published online: 01 Jan 2012.

Option Pricing With Model-Guided Nonparametric Methods

Jianqing FAN and Loriano MANCINI

Parametric option pricing models are widely used in finance. These models capture several features of asset price dynamics; however, their pricing performance can be significantly enhanced when they are combined with nonparametric learning approaches that learn and correct empirically the pricing errors. In this article we propose a new nonparametric method for pricing derivatives assets. Our method relies on the state price distribution instead of the state price density, because the former is easier to estimate nonparametrically than the latter. A parametric model is used as an initial estimate of the state price distribution. Then the pricing errors induced by the parametric model are fitted nonparametrically. This model-guided method, called automatic correction of errors (ACE), estimates the state price distribution nonparametrically. The method is easy to implement and can be combined with any model-based pricing formula to correct the systematic biases of pricing errors. We also develop a nonparametric test based on the generalized likelihood ratio to document the efficacy of the ACE method. Empirical studies based on S&P 500 index options show that our method outperforms several competing pricing models in terms of predictive and hedging abilities.

KEY WORDS: Generalized likelihood ratio test; Model misspecification; Nonparametric regression; Out-of-sample analysis; State price distribution.

1. INTRODUCTION

The last three decades have seen substantial efforts in extending the Black and Scholes (1973) model along several directions. These efforts have aimed to develop more flexible dynamics of asset prices, leading to more accurate option pricing formulas. Examples include the jump-diffusion models of Bates (1991) and Madan, Carr, and Chang (1998), the stochastic volatility models of Hull and White (1987) and Heston (1993), the stochastic volatility and stochastic interest rates models of Amin and Ng (1993) and Bakshi and Chen (1997), and the stochastic volatility jump-diffusion models of Bates (1996) and Scott (1997), among others. These models have substantially relaxed the restrictions in the seminal work of Black and Scholes and made the assumptions of the price movements more plausible. But these models are not derived from comprehensive economic theories, often rely on different assumptions concerning the risk neutral asset dynamics, and need to be simple and convenient to allow for the derivation of pricing formulas. Thus these models cannot be expected to capture all of the relevant features of the pricing mechanisms involved. In fact, there are always limitations on the performance of parametric modeling techniques, and model misspecification is a major concern that can lead to erroneous valuations and hedging strategies.

In this article, instead of attempting to improve option pricing models by introducing even more flexible models, we propose a way to achieve such improvement in an orthogonal direction. Our approach is based on the nonparametric correction of pricing errors induced by a given parametric model. We calibrate the chosen parametric model to best fit the observed option prices, then estimate and correct nonparametrically the pricing errors induced by the given model. Precisely, we estimate the state price survivor function (i.e., 1 minus the state price distribution), using a model-guided nonparametric procedure. We do this because European option prices can be easily expressed in terms of the state price survivor function. This function is always decreasing, has a waterfall shape, and is in a neighborhood of a model-based distribution. We exploit this prior knowledge by estimating the main shape of the survivor function using a sensible parametric model. We then correct the pricing errors induced by the model using a nonparametric approach. This method, called the automatic correction of errors (ACE) of a pricing formula, can be combined with any model-based pricing formula and significantly reduces pricing errors. It does not depend sensitively on the initial model-based calibration, because the nonparametric method in the second step corrects the modeling biases in the initial step. In addition, compared with direct nonparametric methods, the guidance of model-based pricing formulas provides reliable initial option prices and exploits the rich developments in option pricing modeling, resulting in more accurate pricing methods.

Overall, our method departs from the existing option pricing literature in two directions. Most existing studies focus on using the state price density to price options. This density can be estimated nonparametrically by taking the second derivative of call prices, but this procedure is numerically challenging. We rely on the state price distribution, which is related to the first derivative of call prices and thus is easier to estimate. We then nonparametrically fit pricing errors, or deviations from parametric option prices, instead of fitting option prices directly.

We use a nonparametric method to correct the pricing errors, because functional forms of pricing errors are difficult to deter-

Jianqing Fan is Frederick L. Moore'18 Professor of Finance, Department of Operations Research and Financial Engineering, Princeton University, Princeton, NJ 08544, and Honorary Professor, Department of Statistics, Shanghai University of Finance and Economics, Shanghai, China (E-mail: *jqfan@princeton.edu*). Loriano Mancini is Assistant Professor, Swiss Finance Institute at EPFL, CH-1015 Lausanne, Switzerland (E-mail: *loriano.mancini@epfl.ch*). Fan's research was supported by the National Science Foundation grants DMS-0532370 and DMS-0704337. Mancini's research was supported by the University Research Priority Program "Finance and Financial Markets" (University of Zurich) and by the NCCR-FinRisk (Swiss National Science Foundation). This research was undertaken while Mancini visited the Department of Operations Research and Financial Engineering, Princeton University, whose hospitality is gratefully acknowledged. The authors thank the editor, the associate editor, three anonymous referees, Giovanni Barone-Adesi, Eric Ghysels, Rajna Gibson, Peter Gruber, and Claudia Ravanelli for their helpful comments.

Journal of the American Statistical Association
December 2009, Vol. 104, No. 488, Applications and Case Studies
DOI: 10.1198/jasa.2009.ap08171

mine as they vary over time and time to maturity. Nonparametric methods have the flexibility to discover the nonlinear relation between pricing errors and moneyness, that is, the strike price over the forward price of the asset. In the nonparametric literature (e.g., Fan and Yao 2003), it is well known that survivor functions are easier to estimate than density functions, admitting a faster rate of convergence. Thus we estimate the state price survivor function instead of the state price density. This is another important aspect of our methodological contribution to option pricing. The nonparametric learning and correction of pricing errors are easy to implement and fast to compute. The overall procedure is much faster than calibration-based approaches, such as those of Duan (1995), Bakshi, Cao, and Chen (1997), Heston and Nandi (2000), and Barone-Adesi, Engle, and Mancini (2008).

In our empirical analysis, we consider European options on the S&P 500 index from January 2002 to December 2004. We compare our ACE method with (a) the benchmark ad hoc Black–Scholes model of Dumas, Fleming, and Whaley (1998), (b) the parametric generalized autoregressive conditional heteroscedasticity (GARCH) option pricing model of Heston and Nandi (2000), (c) the semiparametric Black–Scholes model developed in this article and inspired by Aït-Sahalia and Lo (1998), and (d) the nonparametric regression approach applied directly to estimate the state price survivor function. We find that our ACE method outperforms both parametric models (ad hoc Black–Scholes and GARCH models) and semiparametric and nonparametric methods in terms of fitting and prediction of option prices, as well as hedging performance. Finally, we develop a nonparametric test based on the generalized likelihood ratio test of Fan, Zhang, and Zhang (2001) to document the efficacy of the ACE method. The test results show that the nonparametric correction of the ACE approach is very effective in reducing the pricing errors. These empirical results provide stark evidence of the power of nonparametric learning of pricing errors for pricing options.

The article is organized as follows. Section 2 introduces our ACE pricing method, and Section 3 develops the nonparametric validation test. Section 4 recalls the semiparametric Black–Scholes and the GARCH pricing models. Section 5 presents the empirical results. Section 6 concludes.

2. ESTIMATION OF THE STATE PRICE SURVIVOR FUNCTION

The basic idea in our approach is to estimate the state price survivor function using portfolios of traded options. Let S_t be the underlying asset price at time t and $f^*(\cdot)$ be the state price (or risk-neutral) conditional density of S_T given information at time t (e.g., Harrison and Kreps 1979). The dependence of the density $f^*(\cdot)$ on time t, time to maturity $\tau = T - t$, and other parameters are suppressed. Let C_t denote the price of the call option at time t, written on the asset S, with strike price X and time to maturity τ, whose payoff function is $\psi(S_T) = \max(S_T - X, 0)$. Then C_t is the discounted expected payoff in the risk-neutral world,

$$C_t = e^{-r_{t,\tau}\tau} E[\psi(S_T)] = e^{-r_{t,\tau}\tau} \int_X^\infty (y - X) f^*(y)\, dy,$$

where $r_{t,\tau}$ is the risk-free rate at time t for the maturity $T = t + \tau$. Let $F^*(x)$ be the cumulative state price distribution of S_T under the risk-neutral measure, that is, $F^*(x) = \int_0^x f^*(y)\, dy$. Integration by parts yields

$$C_t = e^{-r_{t,\tau}\tau} \int_X^\infty \bar{F}^*(y)\, dy, \tag{1}$$

where $\bar{F}^*(x) = 1 - F^*(x)$ is the state price survivor function of S_T. Equation (1) has an interesting economic interpretation. Suppose that n digital call options are available with strike prices $X + \delta, X + 2\delta, \ldots, X + n\delta$. Each of the digital options pays \$1 if the stock price at time T is higher than its corresponding strike price and 0 otherwise. The forward price (before discounting) of each digital call option is $\bar{F}^*(X + i\delta)$, $i = 1, \ldots, n$. When δ is small and n is large, the portfolio Π of long positions δ in each of the digital call options has nearly the same payoff as the original call option, $\max(S_T - X, 0)$. Thus, in terms of forward prices,

$$E[\Pi] = \sum_{i=1}^n \bar{F}^*(X + i\delta)\delta \approx \int_X^\infty \bar{F}^*(y)\, dy = e^{r_{t,\tau}\tau} C_t.$$

This approximation suggests that the integral of the state price survivor function is indeed a portfolio of digital call options. Recently, Giacomini et al. (2008) exploited Equation (1) to price options with mixtures of t-distributions. In general, the price of any derivative contract with payoff function $\psi(S_T)$ can be easily expressed in terms of the state price survivor function,

$$\int_0^\infty \psi(y) f^*(y)\, dy = \psi(0) + \int_0^\infty \bar{F}^*(y)\psi'(y)\, dy,$$

where ψ' is the first derivative of ψ. The payoff function, ψ, must satisfy some mild regularity conditions that are usually verified by derivative contracts traded on the market. The function ψ must be bounded at 0 [i.e., $\psi(0) < \infty$] and not increase too rapidly [i.e., $\lim_{y\to+\infty} \psi(y)\bar{F}^*(y) = 0$], and the integral on the right side must be well defined. Thus for pricing purposes, knowing the state price survivor function is equivalent to knowing the state price density. Of course, if the final aim is to price call options only, then the call pricing function should be modeled directly. Our approach also can be adapted to this application by using model-guided nonparametric techniques to estimate the call pricing function, which will improve pricing accuracy. However, our method based on the state price survivor function is more general and also can be used to price other derivative contracts or less liquid options.

Let $F_{t,\tau} = S_t e^{(r_{t,\tau} - \delta_{t,\tau})\tau}$ be the forward price of the asset at time t and $\delta_{t,\tau}$ be the dividend yield paid by the asset between t and $T = t + \tau$. To avoid any confusion with state price distributions, we always denote the forward price using both subscripts $F_{t,\tau}$. By a change of variables,

$$C_t = e^{-r_{t,\tau}\tau} F_{t,\tau} \int_m^\infty \bar{F}(u)\, du, \tag{2}$$

where $m = X/F_{t,\tau}$ is called the moneyness and $\bar{F}(u) = 1 - F^*(F_{t,\tau}u)$ is the state price survivor function in the normalized scale; that is, $F_{t,\tau}$ is normalized to \$1. For example, the Black and Scholes (1973) model is characterized by the lognormal state price survivor function,

$$\bar{F}(m) = \int_m^\infty \frac{1}{u\sqrt{2\pi\sigma^2\tau}} \exp\left[-\frac{[\log(u)+\sigma^2\tau/2]^2}{2\sigma^2\tau}\right] du$$
$$= 1 - \Phi\left[\frac{\log(m)+\sigma^2\tau/2}{\sigma\sqrt{\tau}}\right], \quad (3)$$

where σ is the constant volatility and Φ is the standard Gaussian distribution function. Under this model, the option price in (2) can be calculated explicitly, resulting in the celebrated Black–Scholes pricing formula,

$$C_t = C_{\mathrm{BS}}(\sigma) = e^{-r_{t,\tau}\tau}(F_{t,\tau}\Phi(d_1) - X\Phi(d_2)), \quad (4)$$

where $d_1 = (\log(F_{t,\tau}/X) + \sigma^2\tau/2)/(\sigma\sqrt{\tau})$ and $d_2 = d_1 - \sigma\sqrt{\tau}$. Equation (4) emphasizes the dependence of the Black–Scholes option price on the volatility, σ. In general, state price distributions have no closed-form expression, and numerical procedures usually are adopted in computations. For example, in jump-diffusion models with nonparametric Lévy measures, Cont and Tankov (2004) suggested approximating the state price density by a mixture of log-normal distributions.

Using Equation (2), pricing European options reduces to estimating the state price survivor function, $\bar{F}$. The parametric approach assumes a risk-neutral dynamic for the underlying asset and then derives the state price survivor function. For example, assuming that S_t follows a geometric Brownian motion implies the state price survivor function in equation (3). Then the parametric approach infers the model parameters from traded options. Our approach is nonparametric. We infer the state price survivor function directly from traded options. The key advantage of this method is that we need neither to assume a parametric model under the risk-neutral measure, nor to derive an analytic form for the pricing formula. This eliminates model misspecification risk and allows fast estimation of the state price survivor function.

2.1 Traded Options and State Price Survivor Function

In this section we discuss how to infer the state price survivor function from traded options. Let $X_1 < X_2$ be two consecutive strike prices of traded options with the same maturity. The portfolio of long positions in $(X_2 - X_1)^{-1}$ call options with strike X_1 and short positions in $(X_2 - X_1)^{-1}$ call options with strike X_2 has a payoff function close to a digital call option with payoff function $I(S_T > (X_1 + X_2)/2)$, where I takes a value of 1 if S_T exceeds $(X_1 + X_2)/2$ and 0 otherwise (Figure 1). Thus

$$e^{r_{t,\tau}\tau}\frac{C_t(X_1) - C_t(X_2)}{X_2 - X_1} \approx E\big[I(S_T > (X_1 + X_2)/2)\big]$$
$$= \bar{F}^*((X_1 + X_2)/2),$$

where $C_t(X_i)$ is the price of the call option with strike X_i at time t. As derived in Appendix A, the midpoint approximation gives the best accuracy in terms of the order of the approximation error. We summarize the theoretical findings in the following proposition.

Proposition 1. Let $C_t(X)$ be the price of the European call option with strike price X at time t. For two consecutive strike prices $X_1 < X_2$, we have

$$e^{r_{t,\tau}\tau}\frac{C_t(X_1) - C_t(X_2)}{X_2 - X_1} = \bar{F}(\bar{m}_{t,1}) + O((m_{t,1} - m_{t,2})^3), \quad (5)$$

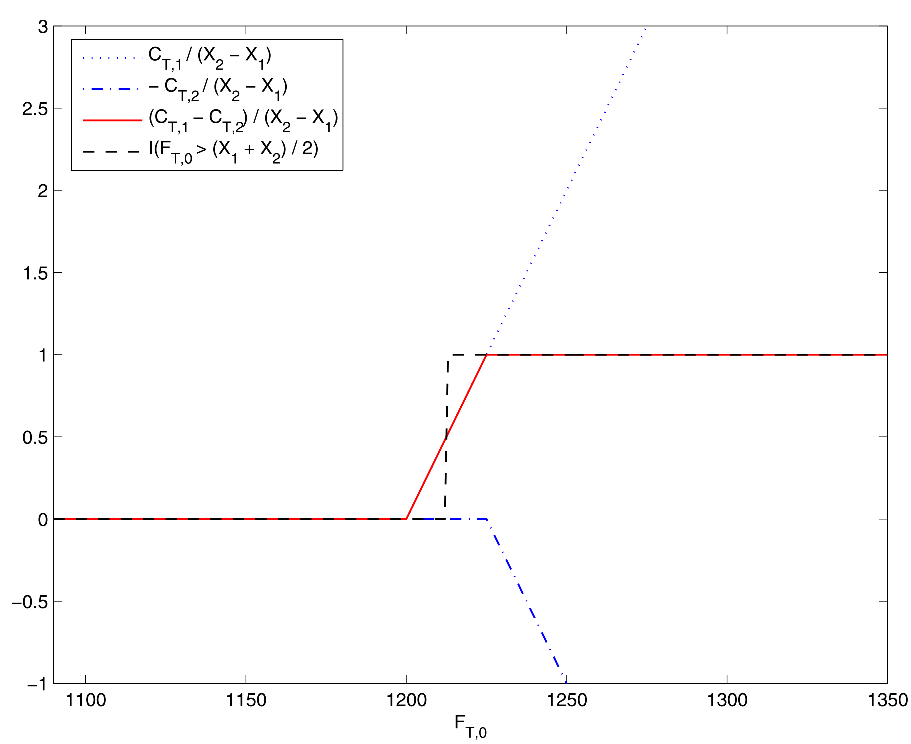

Figure 1. Payoff function (solid line) of the portfolio: long 0.04 shares on the call option with strike price $X_1 = \$1200$ and short 0.04 shares on the call option with strike price $X_2 = \$1225$. This payoff function can be approximated by an indicator function (dashed line). The dotted and dashed–dotted lines are the payoff functions of the long- and short-call options.

where $m_{t,i} = X_i/F_{t,\tau}$, $i = 1, 2$, is the moneyness and $\bar{m}_{t,1} = (m_{t,1} + m_{t,2})/2$. The approximation error is bounded by

$$-\frac{1}{24}\left\{\min_{m_{t,1}\le\xi\le m_{t,2}} f'(\xi)\right\}(m_{t,2} - m_{t,1})^3,$$

where $f'(\xi)$ is the derivative of the normalized state price density, that is, $f'(x) = F''(x)$.

The proof of this proposition is given in Appendix A.

Proposition 1 provides a theoretical basis for inferring state price survivor functions from traded call options. State price distribution can be similarly recovered from the corresponding portfolio of traded put options. Equation (5) suggests that the state price survivor function can be recovered nonparametrically by taking the first derivative of the call price with respect to the strike price. To simplify the notation, let $C_{t,i} = C_t(X_{t,i})$ denote the call option price with moneyness $m_{t,i} = X_i/F_{t,\tau}$ at time t. Order the moneyness $\{m_{t,i}\}$ of traded call options at time t in an ascending order and set

$$\bar{m}_{t,i} = (m_{t,i} + m_{t,i+1})/2 \qquad \text{and} \qquad Y_{t,i} = e^{r_{t,\tau}\tau}\frac{C_{t,i} - C_{t,i+1}}{X_{i+1} - X_i}.$$

Then, according to Equation (5) we have

$$Y_{t,i} = \bar{F}(\bar{m}_{t,i}) + \varepsilon_{t,i}, \tag{6}$$

where $\varepsilon_{t,i}$ is the idiosyncratic noise. This approach reduces the pricing of options to a nonparametric estimation of the function $\bar{F}$, based on the observable data $\{(\bar{m}_{t,i}, Y_{t,i}), i = 1, \ldots, N_t\}$, where $N_t + 1$ is the number of options traded at time t for a given maturity. In contrast to other nonparametric methods, such as those of Aït-Sahalia and Lo (1998) and Aït-Sahalia and Duarte (2003), our method nonparametrically estimates the state price survivor function rather than the state price density. Equation (5) shows that the former, $\bar{F}(m)$, is easily derived from options data. In contrast, the state price density usually is recovered by taking the second derivative of the call option function with respect to the strike price, which is a numerically challenging procedure (see Breeden and Litzenberger 1978). The state price distribution is much easier to estimate, admitting a faster rate of convergence than the state price density (see, e.g., Fan and Yao 2003).

The nonparametric method adopted here, the local linear regression, has several advantages, including automatic boundary correction, high statistical efficiency, and easy bandwidth selection, as noted by Fan (1992) and Fan and Gijbels (1995). (For an overview of the local linear estimator and other related techniques, see Fan and Yao 2003.) On a given day t_0, the nonparametric estimator of $\bar{F}$ is given by the time-weighted local linear regression,

$$\min_{\beta_0,\beta_1\in\mathbb{R}^2} \sum_{t=t_0-d}^{t_0+d} \lambda^{|t_0-t|} \sum_{i=1}^{N_t} (Y_{t,i} - \beta_0 - \beta_1(\bar{m}_{t,i} - m))^2 \times K_h(\bar{m}_{t,i} - m), \tag{7}$$

where $\lambda \in (0, 1]$ is the smoothing parameter in time, K is the kernel function, h is the bandwidth used to fit the local linear model, and $K_h(u) = h^{-1}K(u/h)$. Letting $\hat{\beta}_0$ and $\hat{\beta}_1$ denote the resulting minimizers, $\hat{\bar{F}}(m) = \hat{\beta}_0$ is the nonparametric estimate of the state price survivor function at moneyness m. With the estimated $\hat{\bar{F}}$, for example, call option prices are computed using Equation (2). The first summation in (7) aggregates options data on consecutive dates exploiting the time continuity of the state price survivor function. Without this time aggregation, the sample data available on t_0 are likely insufficient for accurate estimation of $\bar{F}$. In our empirical application, we set $d = 2$ (i.e., 1 trading week), achieving a sample size of 150–200 data points. Options traded on consecutive days have slightly different times to maturity. The time weight, $\lambda^{|t_0-t|}$, accounts for this effect. The second summation in (7) is the standard local linear regression that approximates the function $\bar{F}(x)$ locally around a given point, m, by the linear function

$$\bar{F}(x) \approx \bar{F}(m) + \bar{F}'(m)(x - m) = \beta_0 + \beta_1(x - m)$$

for x in a neighborhood of m. The kernel weights $K_h(\bar{m}_{t,i} - m)$ are used to ensure that the regression is run locally. For instance, using the Epanechnikov kernel, $K(x) = \frac{3}{4}(1-x^2)I(|x| \le 1)$, that vanishes outside the interval $(-1, 1)$, the local linear regression in (7) uses only options data with moneyness $\bar{m}_{t,i}$ in the interval $m \pm h$. Thus the bandwidth controls the effective sample size. For example, Figure 2 shows the nonparametric estimate of the survivor function $\bar{F}$ on December 29, 2004, using $Y_{t,i}$'s observed on December 27–31, 2004, with time maturities ranging from 173 to 169 days. Visual inspection of the fitting suggests that the nonparametric method performs well. Unfortunately, however, small errors in estimating the state price survivor function can translate into large pricing errors. Figure 3 shows that this is the case for the options traded on December 29, 2004.

2.2 Ad hoc Black–Scholes Model

The implied volatility, $\sigma_{t,i}^{\mathrm{BS}} = C_{\mathrm{BS}}^{-1}(C_{t,i})$, is a common measure used to represent the call price, $C_{t,i}$. A well-documented empirical feature of implied volatilities is the so-called "volatility smile" (e.g., Renault and Touzi 1996). For example, Figure 4 shows the implied volatilities of the call option prices analyzed in the previous section and observed on December 27–31, 2004. To account for this phenomenon and provide a benchmark option pricing model, Dumas, Fleming, and Whaley (1998) introduced an ad hoc Black–Scholes model in which the implied volatilities are smoothed across moneyness by fitting a parabolic function,

$$\sigma_{t,i}^{\mathrm{BS}} = a_0 + a_1 m_{t,i} + a_2 m_{t,i}^2 + \mathrm{error}_{t,i}, \tag{8}$$

where $\sigma_{t,i}^{\mathrm{BS}}$ denotes the implied volatility observed on day t for a given maturity T and moneyness $m_{t,i}$, $t \in [t_0 - d, t_0 + d]$ and $i = 1, \ldots, N_t + 1$. The implied volatilities observed on different days have slightly different times to maturity, $\tau = T - t$, for $t \in [t_0 - d, t_0 + d]$. In our empirical application, for each day t_0 and each maturity T, the quadratic function (8) is estimated using the least squares regression with time weight, $\lambda^{|t_0-t|}$, as for nonparametric regression (7). To price an option with moneyness m and time to maturity τ, the fitted value, $\hat{\sigma}(m) = \hat{a}_0 + \hat{a}_1 m + \hat{a}_2 m^2$, is plugged into the Black–Scholes formula (4) to obtain an ad hoc Black–Scholes pricing formula, $C^{\mathrm{BS}}(m) = C_{\mathrm{BS}}(\hat{\sigma}(m))$. Figure 4 shows the quadratic fit to the implied volatility. Some lack of fit is evidenced because of the inflexibility of the quadratic form. This translates into systematic pricing errors, as demonstrated in Figure 3.

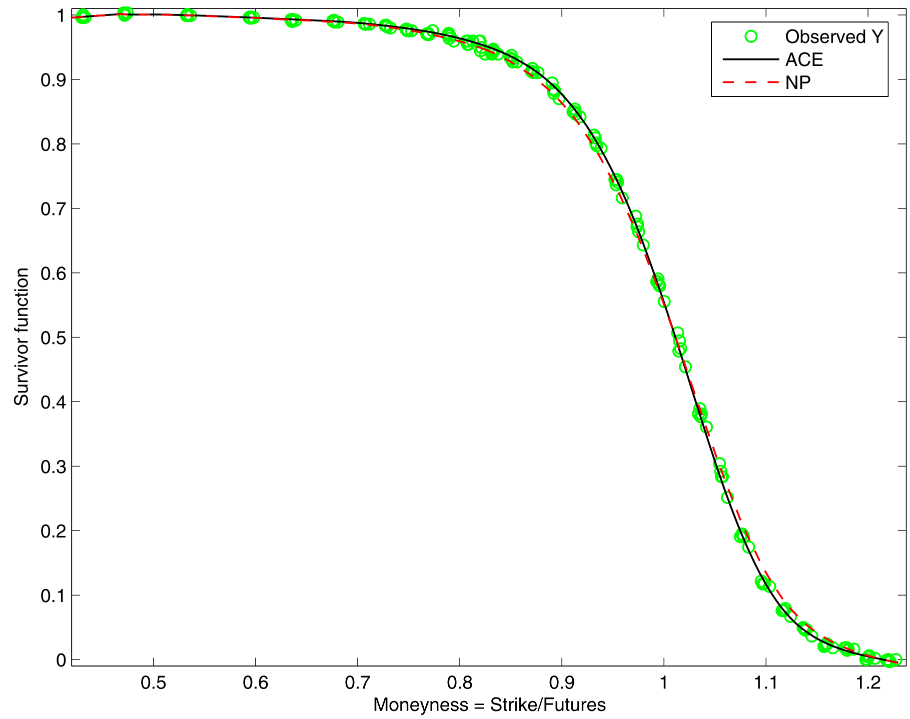

Figure 2. Scatterplot of $Y_{t,i} = e^{r_{t,\tau}\tau}(C_{t,i} - C_{t,i+1})/(X_{i+1} - X_i)$ versus the moneyness, $\bar{m}_{t,i}$, for the call option prices observed on December 27–31, 2004, with maturities of 169–173 days. The survivor function is estimated by using the direct nonparametric approach (7) (NP) and our proposed ACE approach (11).

Although theoretically inconsistent, ad hoc Black–Scholes methods are routinely used in the option pricing industry. They represent a challenging benchmark because they allow for different implied volatilities to price different options, and in our empirical applications they are separately fitted to each maturity. Moreover, Dumas, Fleming, and Whaley (1998) showed

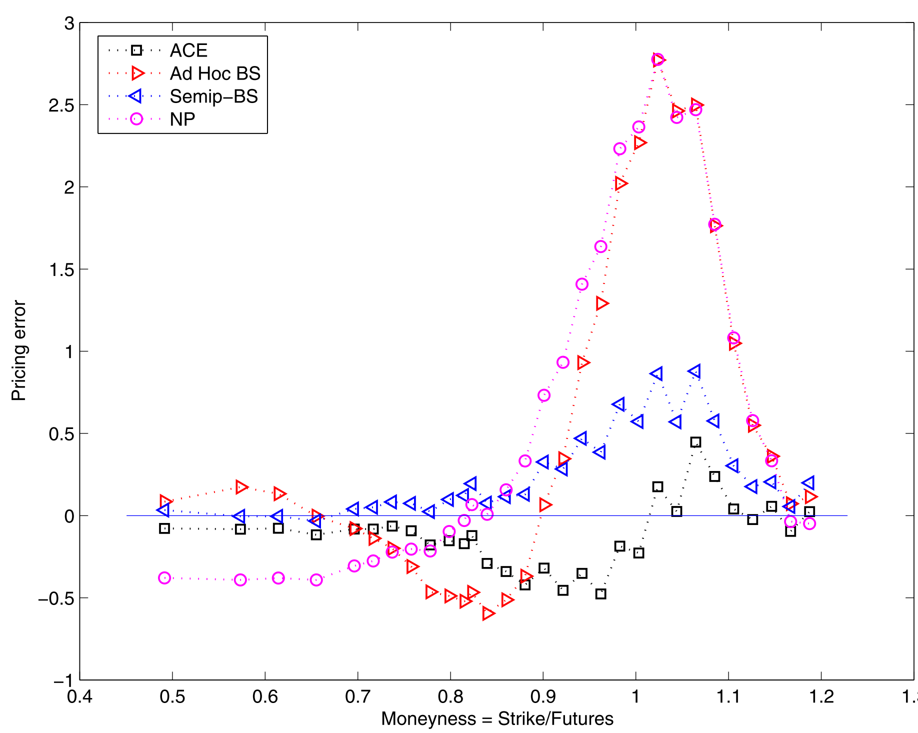

Figure 3. Pricing performance of the direct nonparametric approach (NP), the ad hoc Black–Scholes model (Ad Hoc BS), the ACE approach, and the semiparametric Black–Scholes (Semip-BS) model. The graph shows the dollar pricing error (model price − market price), on December 29, 2004. All methods are estimated using call prices on December 27–31, 2004, with maturities of 169–173 days.

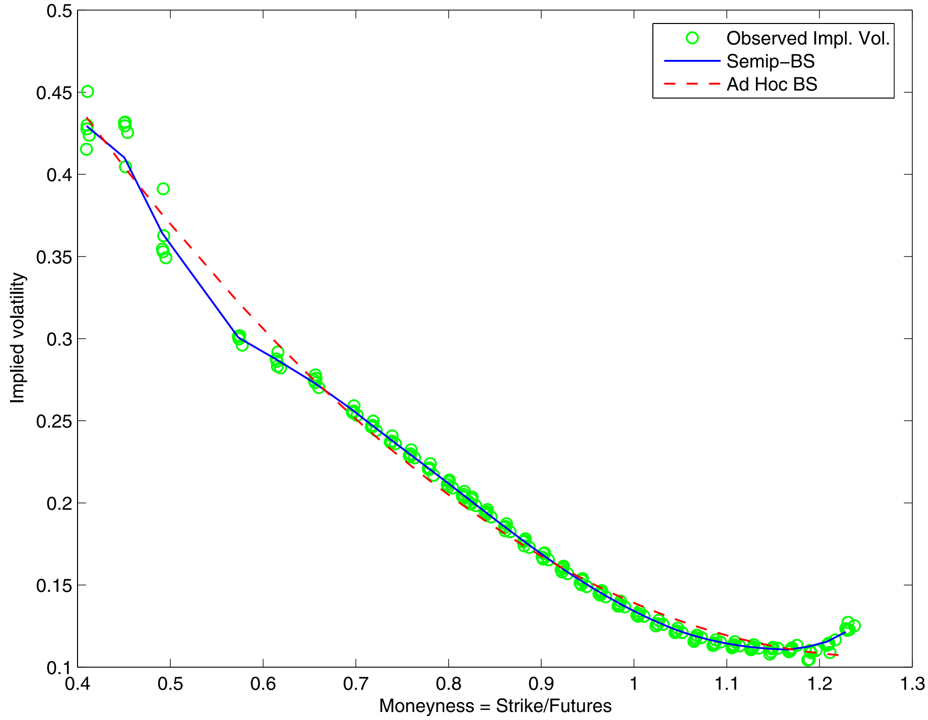

Figure 4. Implied volatilities observed on December 27–31, 2004 for call option prices with maturities of 169–173 days. The figure shows the fitted parabola, $\sigma = a_0 + a_1 m + a_2 m^2$ (dashed line), and the local linear estimation of the function $\sigma(m)$ (solid line), where m denotes the moneyness.

that this model outperforms deterministic volatility function models introduced by Derman and Kani (1994), Dupire (1994), and Rubinstein (1994).

2.3 Model-Guided Nonparametric Methods

The direct nonparametric approach (7) does not perform well, because it does not exploit the prior knowledge on the shape of the state price survivor function, it uses a constant bandwidth h over the whole domain of $\bar{F}$, and it does not explicitly account for the implied volatility smile. We propose addressing these shortcomings using model-guided nonparametric estimation of the state price survivor function, $\bar{F}$.

To estimate the main shape of the survivor function, we combine the state price distribution (3) and the ad hoc Black–Scholes model (8). This results in the following preliminary estimator of the survivor function:

$$\bar{F}_{\mathrm{LN}}(m;\vartheta_t) = 1 - \Phi\left[\frac{\log(m) + (\hat{\sigma}(m)\vartheta_t)^2/2}{\hat{\sigma}(m)\vartheta_t}\right], \tag{9}$$

where $\hat{a}_0$, $\hat{a}_1$, and $\hat{a}_2$ in $\hat{\sigma}(m)$ are estimated using the implied volatilities observed on $2d+1$ consecutive days as in the ad hoc Black–Scholes model, Equation (8). The parameter ϑ_t accounts for the slightly different times to maturity and is determined by minimizing the following distance between empirical and theoretical values:

$$\hat{\vartheta}_t = \arg\min_{\vartheta\in\mathbb{R}} \sum_{i=1}^{N_t} (Y_{t,i} - \bar{F}_{\mathrm{LN}}(\bar{m}_{t,i};\vartheta))^2. \tag{10}$$

Note that $\bar{F}_{\mathrm{LN}}(m;\vartheta_t)$ reduces to the Black–Scholes lognormal survivor function (3) when $a_1 = a_2 = 0$, and $\vartheta_t = \sqrt{T-t}$. The main shape of the survivor curve is now captured by our preliminary estimate $\bar{F}_{\mathrm{LN}}(m;\hat{\vartheta}_t)$. The accuracy of this estimate is not so important, because it will be corrected by a nonparametric estimation in the second stage. Indeed, any state price survivor function $\bar{F}_t(m)$ can be represented as

$$\bar{F}_t(m) = \bar{F}_{\mathrm{LN}}(m;\hat{\vartheta}_t) + \bar{F}_{t,c}(m), \tag{11}$$

where $\bar{F}_{\mathrm{LN}}(m;\hat{\vartheta}_t)$ is the parametric leading term of the state price survivor function and $\bar{F}_{t,c}(m)$ is the nonparametric correction term. Equations (6) and (11) give

$$Y_{t,i} = \bar{F}_{\mathrm{LN}}(\bar{m}_{t,i};\hat{\vartheta}_t) + \bar{F}_{t,c}(\bar{m}_{t,i}) + \varepsilon_{t,i}. \tag{12}$$

This equation shows that $\bar{F}_{t,c}(m)$ is the regression function for the data $\tilde{Y}_{t,i} = Y_{t,i} - \bar{F}_{\mathrm{LN}}(\bar{m}_{t,i};\hat{\vartheta}_t)$ on the moneyness $\bar{m}_{t,i}$. The $\tilde{Y}_{t,i}$'s are the pricing errors on the digital call options induced by the parametric model $\bar{F}_{\mathrm{LN}}(m;\hat{\vartheta}_t)$. These pricing errors are estimated and corrected by the nonparametric correction term $\bar{F}_{t,c}(m)$. This term can be estimated by the local linear fit to the data, $\{(\bar{m}_{t,i}, \tilde{Y}_{t,i}), i = 1,\dots,N_t; t \in [t_0-d, t_0+d]\}$, using the time-weighted nonparametric regression,

$$\min_{\beta_0,\beta_1\in\mathbb{R}^2} \sum_{t=t_0-d}^{t_0+d} \lambda^{|t_0-t|} \sum_{i=1}^{N_t} (\tilde{Y}_{t,i} - \beta_0 - \beta_1(\bar{m}_{t,i} - m))^2 \times K_h(\bar{m}_{t,i} - m). \tag{13}$$

Letting $\tilde{\beta}_0$ and $\tilde{\beta}_1$ denote the resulting minimizers, the nonparametric correction term at the moneyness m is $\hat{\bar{F}}_{t,c}(m) = \tilde{\beta}_0$. We applied nonparametric regression to the $\tilde{Y}_{t,i}$'s. These are more homogeneous than the $Y_{t,i}$'s, with approximately the same degree of smoothness as a function of m; thus the constant bandwidth h should be a reasonable choice for estimating the correction term $\bar{F}_{t,c}(m)$. The shape of $\bar{F}_{t,c}(m)$ can vary over time and time to maturity, and here a nonparametric learning technique is particularly appealing, and also easy to implement. This procedure gives a model-guided nonparametric estimation of $\bar{F}_t(m)$ by plugging the nonparametric estimate of the function $\bar{F}_{t,c}(m)$ into (11). Nonparametric estimation of $\bar{F}_{t,c}(m)$ is equivalent to the nonparametric estimation of $\bar{F}_t(m)$; however, the former is easier to estimate nonparametrically, because it is a smoother function of m. We note that $\bar{F}_{t,c}$ is not a survivor function, but rather the correction term of the parametric function $\bar{F}_{\mathrm{LN}}$. According to (2), the price of a call option can be summarized as follows.

Proposition 2. The price of the European call option at time t with moneyness m and time to maturity τ can be decomposed as

$$C_t = e^{-r_{t,\tau}\tau} F_{t,\tau} \int_m^\infty \bar{F}_{\mathrm{LN}}(u; \hat{\vartheta}_t)\, du + e^{-r_{t,\tau}\tau} F_{t,\tau} \int_m^\infty \bar{F}_{t,c}(u)\, du. \quad (14)$$

The proof is given by simply substituting Equation (11) into Equation (2). Substituting $\hat{\hat{\bar{F}}}_{t,c}(m)$ into (14), we obtain a new method for pricing derivatives instruments. The last term in (14) is the nonparametric correction of the pricing error induced by the parametric pricing formula. The overall procedure is still nonparametric and can be combined with any parametric approach. We call this method the automatic correction of errors (ACE) approach, where "automatic" refers to the nonparametric fitting, which need not impose any functional form.

As an example, Figure 2 shows the estimated state price survivor function on December 29, 2004 using the same call options as in previous sections and applying our ACE method (11) and the direct nonparametric method (7). Visually, both methods seem to provide a good fit to the data; however, Figure 3 shows that the the tweo methods' pricing errors behave very differently. The direct nonparametric approach does not perform well, with a root mean squared error (RMSE) of \$1.04. Our ACE approach reduces the pricing errors substantially and has an RMSE of only \$0.21. For completeness, Figure 3 also presents the pricing errors of the ad hoc Black–Scholes model (8). This method does not perform well, with a RMSE of \$1.10, mainly because the fitted volatilities $\hat{\sigma}^{\mathrm{BS}}$ in Equation (8) are not very accurate around the moneyness, $m \approx 1$. At-the-money options are very sensitive to changes in volatilities, and even small errors in the volatility estimation can induce large pricing errors. Figure 3 also shows that the pricing performance of the semiparametric Black–Scholes model, introduced in Section 4.1, it is quite satisfactory, with a RMSE of \$0.35.

2.4 Statistical Properties of the Automatic Correction of Errors Method and Bandwidth Selection

In this section we show that the parametrically guided nonparametric fitting of the ACE method has smaller bias when the true state price survivor function is in a neighborhood of the parametric model $\bar{F}(m;\theta)$. The least squares calibration method gives the θ that minimizes

$$\sum_{t=t_0-d}^{t_0+d} \sum_{i=1}^{N_t} (Y_{t,i} - \bar{F}(\bar{m}_{t,i};\theta))^2. \quad (15)$$

Let $n = \sum_{t=t_0-d}^{t_0+d} N_t$ be the sample size. To simplify the presentation here, we assume that the true survivor function, $\bar{F}_0(m)$, is the same from $t_0 - d$ to $t_0 + d$ or varies very slowly in short time periods. If this assumption is violated, then we need to consider the date t_0 only, that is, $d = 0$. The nonlinear least squares method (15) attempts to find θ_0 that minimizes

$$E[\bar{F}_0(m) - \bar{F}(m;\theta)]^2. \quad (16)$$

The survivor function $\bar{F}(m;\theta_0)$ is the best approximation of the true state price survivor function, $\bar{F}_0(m)$, in the family of functions $\{\bar{F}(m;\theta)\}$. The next proposition summarizes the bias and variance of the estimator $\hat{\hat{\bar{F}}}(m) = \bar{F}(m;\hat{\theta}) + \hat{\hat{\bar{F}}}_c(m)$, where $\hat{\hat{\bar{F}}}_c(m)$ is the local linear fit to the data $\{(\bar{m}_{t,i}, \tilde{Y}_{t,i})\}$.

Proposition 3. Under the conditions given in Appendix B, we have

$$\sqrt{nh}\left\{\hat{\hat{\bar{F}}}(m) - \bar{F}_0(m) - \frac{1}{2}\bar{F}_c''(m)h^2 \int u^2 K(u)\,du - o(h^2)\right\} \xrightarrow{w} \mathrm{N}\left(0, \sigma_\varepsilon^2(m) \int K^2(u)\,du / g(m)\right),$$

where $\bar{F}_c(m) = \bar{F}_0(m) - \bar{F}(m;\theta_0)$, $g(m)$ is the marginal density of the moneyness at the point m, and $\sigma_\varepsilon^2(m)$ is the conditional variance of $\varepsilon_{t,i}$ given $\bar{m}_{t,i} = m$.

The proof of this proposition is given in Appendix B.

The bias of the parametric-guided nonparametric estimator (i.e., the ACE method) has a leading term of order $\frac{1}{2}\bar{F}_c''(m)h^2 \times \int u^2 K(u)\,du$, whereas the direct nonparametric estimator has bias $\frac{1}{2}\bar{F}''(m)h^2 \int u^2 K(u)\,du$. The former is much smaller than the latter when $\bar{F}_c$ is smooth and small; this is particularly so when $\bar{F}(m;\theta_0)$ is close to the true state price survivor function. The advantages of a parametrically guided nonparametric regression over the direct nonparametric approach have been documented by Glad (1998) and Fan and Ullah (1999). In our application, the curvature of $\bar{F}_0(m)$ is large when m is around 1. Exploiting the shape of the function $\bar{F}(m;\theta_0)$, the curvature of $\bar{F}_c(m) = \bar{F}_0(m) - \bar{F}(m;\theta_0)$ can be reduced significantly; thus the ACE method performs better than the direct nonparametric approach.

Hjort and Glad (1995) proposed a method similar to ACE to estimate density functions. But their goal was to estimate densities (not distributions) based on a random sample, and they used a multiplicative decomposition of the unknown density function given by a parametric density times a nonparametric correction term, rather than an additive decomposition as in Equation (11). Although our interest lies in estimating the state price distribution, our problem is indeed reduced to a nonparametric regression problem inferred from option prices with different moneyness. We also could use a multiplicative decomposition of the state price survivor function in (11), but this would induce a less straightforward derivation of the adequacy test in Section 3.

Garcia and Gençay (2000) introduced a somewhat similar approach to price options, deriving a generalized Black–Scholes formula but calibrated using neural networks. In the statistics literature, Press and Tukey (1956) and Glad (1998), among others, have used various parametric-guided approaches.

Next, we briefly discuss the issue of the bandwidth selection. Because the problem (11) is a standard nonparametric regression problem, a wealth of data-driven bandwidths can be used (see Fan and Yao 2003). In particular, one can apply the preasymptotic substitution method of Fan and Gijbels (1995) or the plug-in method of Ruppert, Sheather, and Wand (1995). Alternatively, one can choose the bandwidth either subjectively or by a simple rule of thumb. The latter method carries nearly no computational cost, and the selected bandwidth tends to be stable from one day to another, which is particularly important in practical applications. In our empirical study, we simply take $h = 0.3s$, where s is the sample standard deviation of the moneyness $\{\bar{m}_{t,i}, i = 1, \ldots, N_t; t \in [t_0 - d, t_0 + d]\}$. The standard deviation accounts for the dispersion of the moneyness, and the constant factor 0.3 is an empirical choice from trial and error. Other constant factors might be required for other data sets. The ACE method requires only univariate nonparametric estimations and thus can be easily applied using various bandwidths, selecting the bandwidth that induces the best pricing performance.

3. ADEQUACY OF THE PRICING FORMULA

Parametric pricing models are based on assumptions regarding the risk-neutral dynamic of the underlying asset. Thus an important question is whether or not these assumptions are consistent with the observed option prices—in other words, does the pricing model induced by the parametric state price survivor function fit traded options adequately? Statistically, this is a nonparametric hypothesis testing problem,

$$\begin{aligned} &H_0 : \bar{F}_t(m) = \bar{F}_t(m;\theta) \quad \text{versus} \\ &H_1 : \bar{F}_t(m) \neq \bar{F}_t(m;\theta), \end{aligned} \tag{17}$$

based on the observed data from model (6), where $\bar{F}_t(m)$ is the true state price survivor function and $\bar{F}_t(m;\theta)$ is the state price survivor function derived from the parametric model. The null hypothesis is parametric, but the alternative hypothesis is nonparametric. Thus the classical likelihood ratio test needs to be properly extended to deal with such a general situation. One such extension is the generalized likelihood ratio (GLR) test proposed by Fan, Zhang, and Zhang (2001) (see Fan and Jiang 2007 for an overview). In the current setting, the GLR test compares the residual sum of squares when fitting model (6) using the parametric and ACE methods. A direct application of the GLR statistic is not ideal here, however. Even when the null hypothesis is correct, the nonparametric fits incur biases (see Sec. 2.4). To improve the testing procedure, Fan and Yao (2003, chapter 9) suggested testing whether or not the correction term $\bar{F}_{t,c}(m)$ is statistically different from 0. Under the null hypothesis, $\bar{F}_{t,c}$ is 0 or, equivalently, $\bar{F}_t(m) = \bar{F}_t(m;\theta)$, and the residual sum of squares is

$$\mathrm{RSS}_0 = \sum_{t=t_0-d}^{t_0+d} \sum_{i=1}^{N_t} \tilde{Y}_{t,i}^2 I(a \le \bar{m}_{t,i} \le b)$$

for a given large interval $[a,b]$. The survivor function is tested on the interval $[a,b]$, and the parameter θ characterizing $\bar{F}_t(m;\theta)$ is calibrated to traded options on the dates $[t_0 - d, t_0 + d]$ with moneyness falling in $[a,b]$. This procedure ensures that the test results are not driven by potential difficulties of the nonparametric approach in fitting the tails of the survivor functions. Under the alternative hypothesis, $\bar{F}_{t,c}$ is not 0, and the residual sum of squares is

$$\mathrm{RSS}_1 = \sum_{t=t_0-d}^{t_0+d} \sum_{i=1}^{N_t} (\tilde{Y}_{t,i} - \hat{\bar{F}}_{t,c}(\bar{m}_{t,i}))^2 I(a \le \bar{m}_{t,i} \le b).$$

Let $n_{a,b} = \sum_{t=t_0-d}^{t_0+d} \sum_{i=1}^{N_t} I(a \le \bar{m}_{t,i} \le b)$ be the number of data points used in the fitting. The GLR test statistic, defined as

$$T_n = \frac{n_{a,b}}{2} \log(\mathrm{RSS}_0/\mathrm{RSS}_1), \tag{18}$$

measures the inadequacy of the parametric fit. The larger the test statistics T_n, the less adequate the fit of the parametric model to the options data. When T_n is very large (or beyond the usual high quantiles of its asymptotic null distribution), the null hypothesis that $\bar{F}_{t,c}$ is 0 must be rejected. The following proposition derives the asymptotic null distribution. We set $\lambda = 1$ for notational simplicity.

Proposition 4. Under the conditions given in Appendix C, if the null hypothesis is true, then

$$r_K T_n \overset{a}{\sim} \chi^2_{a_n} \tag{19}$$

in the sense that

$$\frac{r_K T_n - a_n}{\sqrt{2a_n}} \xrightarrow{w} \mathrm{N}(0,1),$$

where, with $*$ denoting the convolution operator,

$$\begin{aligned} r_K &= \frac{K(0) - \int K^2(t)\,dt/2}{\int (K(t) - K * K(t)/2)^2\,dt}, \\ a_n &= s_K(b-a)/h + 1.45, \qquad \text{and} \\ s_K &= \frac{(K(0) - \int K^2(t)\,dt/2)^2}{\int (K(t) - K * K(t)/2)^2\,dt}. \end{aligned}$$

The constants r_K and s_K have been computed by Fan, Zhang, and Zhang (2001). For the Epanechnikov kernel, $r_K = 2.1153$ and $s_K = 0.9519$. The constant 1.45 in a_n comes from the empirical formula of Zhang (2003), who also demonstrated the adequacy of such an approximation.

In the foregoing formulation, the parametric model $\bar{F}_t(m;\theta)$ can be any survivor function. In our empirical application, we take the lognormal survivor function (9) under the ad hoc Black–Scholes model (8). We apply the GLR test statistic (18) to the S&P 500 index options from January 2, 2002 to December 31, 2004, as detailed in Section 5. We set the coefficients a and b equal to the 0.05 and 0.95 quantiles of the observed moneyness for each relevant maturity. This procedure ensures that the estimate of the correction term is based on a sufficiently large sample size. Using (19) to compute the p-value, we find that all test statistics have a p-value no larger than 0.001 for every maturity in this 3-year period. As a robustness check, we computed p-values also using the conditional nonparametric bootstrap method of Fan and Yao (2003, chapter 9), to better approximate the null distribution of the GLR test statistic. In nearly all tests, p-values were similar and very low. These results provide stark evidence that the nonparametric correction term, $\bar{F}_{t,c}$, is very effective in reducing the pricing errors of the ad hoc Black–Scholes model, $\bar{F}_{\mathrm{LN}}(m;\vartheta)$.

4. OTHER PRICING METHODS

In an empirical study, we compared our ACE method with two option pricing models.

4.1 Semiparametric Black–Scholes Model

As shown in Figure 4, the parabola in the ad hoc Black–Scholes model might not be sufficiently flexible to fit the implied volatility smile. One way to overcome this difficulty is to fit the implied volatility function nonparametrically. This approach allows for a more flexible functional dependence of the implied volatility on moneyness, $\sigma^{\mathrm{BS}}(m)$. For each day t_0 and maturity T, we use the local linear regression to estimate the implied volatility function directly. As in the other model estimations, we aggregate options data around day t_0. The local linear estimate, $\hat{\sigma}^{\mathrm{BS}}(m)=\hat{\beta}_0$, is given by the time-weighted regression,

$$\min_{\beta_0,\beta_1\in\mathbb{R}^2}\sum_{t=t_0-d}^{t_0+d}\lambda^{|t_0-t|}\sum_{i=1}^{N_t}(\sigma_{t,i}^{\mathrm{BS}}-\beta_0-\beta_1(m_{t,i}-m))^2 \times K_h(m_{t,i}-m), \quad (20)$$

and $\hat{\beta}_0$, $\hat{\beta}_1$ are the resulting minimizers; see (7) for definitions of smoothing parameters λ, bandwidth h, and kernel function K. As in the ad hoc Black–Scholes model (8), here call option prices are computed by plugging $\hat{\sigma}^{\mathrm{BS}}(m)$ into the Black–Scholes formula, $C_{t,i}^{\mathrm{BS}}$, and setting $\sigma=\hat{\sigma}^{\mathrm{BS}}(m)$. This pricing method is inspired by the work of Aït-Sahalia and Lo (1998), who fit two-dimensional functions to the implied volatilities using a different nonparametric functional form. Figure 4 shows the nonparametric fit of the implied volatilities on December 29, 2004, using the same options as before. The data demonstrate the flexibility of the nonparametric fitting.

4.2 Generalized Autoregressive Conditional Heteroscedasticity Option Pricing Model

Financial asset returns exhibit variances that change over time (e.g., Schwert 1989; Jones 2003). The time-varying volatility of asset returns often is described using GARCH models (see Engle 1982; Bollerslev 1986; Bollerslev, Chou, and Kroner 1992; Ghysels, Harvey, and Renault 1996). In our empirical analysis we consider the parametric GARCH option pricing of Heston and Nandi (2000), who derived an almost closed-form pricing formula. Here we describe only the main features of the model; a detailed description has been provided by Heston and Nandi (2000). Under the risk-neutral distribution, the log-price follows the following GARCH model:

$$\begin{aligned}\log(S_t/S_{t-1})&=r-h_t/2+\sqrt{h_t}z_t,\\ h_t&=\omega+\beta h_{t-1}+\alpha(z_{t-1}-\gamma\sqrt{h_{t-1}})^2,\end{aligned}\quad (21)$$

where z_t is a Gaussian innovation and h_t is the conditional variance of the log-return between $t-1$ and t, given the information, $\mathcal{I}_{t-1}$, available at time $t-1$. When $\beta+\alpha\gamma^2<1$, the log-return process is stationary with finite mean and variance. The parameters α and γ determine the kurtosis and the asymmetry of the distribution. When $\gamma>0$, the model accounts for the so-called "leverage effect"; that is, a negative shock, z_t, increases the variance more than a positive shock, z_t, of the same absolute magnitude. At time $t=0$, the price of the call option with strike price X and maturity T is

$$\begin{aligned}C^{\mathrm{HN}}&=e^{-rT}E[\max(S_T-X,0)]\\ &=e^{-rT}\zeta(1)\left(\frac{1}{2}+\frac{1}{\pi}\int_0^\infty\Re\left[\frac{X^{-i\phi}\zeta(i\phi+1)}{i\phi\zeta(1)}\right]d\phi\right)\\ &\quad-e^{-rT}X\left(\frac{1}{2}+\frac{1}{\pi}\int_0^\infty\Re\left[\frac{X^{-i\phi}\zeta(i\phi)}{i\phi}\right]d\phi\right),\end{aligned}\quad (22)$$

where $\Re[\cdot]$ denotes the real part of a complex number, $i=\sqrt{-1}$, and $\zeta(\phi)$ is the moment-generating function at time t of the log-price, $p_T=\log(S_T)$,

$$\zeta(\phi)=E[e^{\phi p_T}|\mathcal{I}_t]=e^{\phi p_t+A_t+B_t h_{t+1}}.$$

The coefficients A_t and B_t are computed backward using recursive equations (see Heston and Nandi 2000; Christoffersen, Heston, and Jacobs 2006). Given the past underlying returns, the current variance, h_{t+1}, is known at time t. This is an important advantage of GARCH pricing models over other stochastic volatility models, such as those of Heston (1993) and Bakshi, Cao, and Chen (1997), in which the instantaneous variance is not observed and requires calibration or separate estimation.

5. EMPIRICAL ANALYSIS

5.1 The Data

We consider closing prices of European options on the S&P 500 index (symbol SPX) from January 2, 2002 to December 31, 2004. The data are downloaded from OptionMetrics. The average of bid and ask prices is taken as the option price. Such mid prices are not transaction prices (which are not available in OptionMetrics), but they are all recorded at the same time, which simplifies the analysis. To retain only liquid options (e.g., Aït-Sahalia and Lo 1998), options with implied volatility $>70\%$, prices $\le 1/8$, and times to maturity (in calendar days) <20 days or >240 days are discarded, yielding a sample of 101,036 observations.

The market for SPX options is one of the most active index options market worldwide. Expiration months are the three near-term months and three additional months from the March–June–September–December quarterly cycle. Strike price intervals are 5 and 25 points. The options are European, have no wild card features, and can be hedged using the active market on the S&P 500 index futures. Consequently, SPX options have been the focus of many empirical investigations, including those of Aït-Sahalia and Lo (1998), Chernov and Ghysels (2000), Heston and Nandi (2000), Carr et al. (2003), and Barone-Adesi, Engle, and Mancini (2008).

The term structure of default-free interest rates is also downloaded from OptionMetrics, and the riskless interest rate for each maturity, τ_i, is obtained by linearly interpolating the two interest rates whose maturities straddle τ_i. This procedure is repeated for each contract and each day in the sample.

Table 1. Database description

		Maturity					
		Less than 60		60 to 160		More than 160	
		Mean	SD	Mean	SD	Mean	SD
DITM	Call price $	323.22	89.16	342.00	108.12	354.06	117.30
	σ^{BS} (%)	43.91	9.70	35.67	7.82	30.57	5.40
	Observations	6,280		9,115		4,314	
ITM	Call price $	129.49	41.32	141.22	38.74	153.38	36.11
	σ^{BS} (%)	26.99	6.84	24.79	5.44	23.19	4.21
	Observations	7,969		8,609		3,840	
ATM	Call price $	30.61	18.60	45.47	19.45	64.27	19.20
	σ^{BS} (%)	18.06	5.82	19.16	5.01	19.36	4.08
	Observations	10,832		8,058		3,028	
OTM	Call price $	3.03	4.02	7.88	7.76	17.28	11.63
	σ^{BS} (%)	18.86	5.79	17.20	4.74	17.13	4.04
	Observations	7,395		9,175		4,302	
DOTM	Call price $	0.28	0.16	0.51	0.78	1.39	2.15
	σ^{BS} (%)	37.38	11.55	25.91	8.40	20.32	4.72
	Observations	4,561		8,775		4,783	

NOTE: Shown are the mean, standard deviation (SD), and number of observations for each moneyness and maturity category of SPX call option prices from January 2, 2002 to December 31, 2004, after applying filtering criteria and replacing illiquid ITM options as described in the text. Here σ^{BS} is the Black–Scholes implied volatility, DITM is deep in-the-money options with moneyness < 0.8, ITM is in-the-money options with moneyness between 0.8 and 0.94, ATM is at-the-money options with moneyness between 0.94 and 1.04, OTM is out-of-the-money options with moneyness between 1.04 and 1.2, and DOTM is deep out-of-the-money options with moneyness > 1.2. Moneyness is defined as strike price divided by the forward price of the underlying asset. Maturity is measured in calendar days.

The raw data present three challenges. First, in-the-money (ITM) options are not actively traded compared with at-the-money (ATM) and out-of-the-money (OTM) options. For example, even since the October 1987 crash, the daily volume for OTM puts has been usually several times as large as the volume for ITM puts, reflecting the strong demand by portfolio managers for protective puts. Second, it is difficult to observe the underlying index price exactly when option prices are recorded. Temporal mismatches between option and index price recordings can induce pricing biases (e.g., Fleming, Ostdiek, and Whaley 1996). Third, the stocks in the S&P 500 index pay dividends, and the future rates of dividend are difficult to determine. We addressed these three issues following the approach suggested by Aït-Sahalia and Lo (1998). Because option prices are recorded at the same time each day, only one temporally matched index price per day is required for estimating a pricing model. Because the dividend yield, $\delta_{t,\tau}$, is not observable, for each maturity, τ, the forward price, $F_{t,\tau}$, is computed via the put-call parity, which holds because of the absence of arbitrage opportunities, independent of any option pricing model,

$$C_t + Xe^{-r_{t,\tau}\tau} = P_t + F_{t,\tau}e^{-r_{t,\tau}\tau},$$

where P_t denotes the put price. The forward price is computed using liquid calls and puts closest to ATM. The procedure is repeated for all dates and time to maturities. Then the prices of illiquid ITM calls are replaced by the corresponding prices implied by the put-call parity, $P_t + F_{t,\tau}e^{-r_{t,\tau}\tau} - Xe^{-r_{t,\tau}\tau}$, where the put price is out-of-the-money and thus liquid. After this procedure, the information in liquid OTM put prices translates into implied ITM call prices via the put-call parity. Thus put prices may be discarded with no loss of additional information. Some previous empirical studies have investigated model pricing performance using calls and puts separately, and have yielded rather similar findings (e.g., Bakshi, Cao, and Chen 1997; Dumas, Fleming, and Whaley 1998). We divide the call option data into several categories according to either moneyness or time to maturity. A call option is said to be deep in the money (DITM) if its moneyness $m < 0.8$, ITM if $0.8 \le m < 0.94$, ATM if $0.94 \le m < 1.04$, OTM if $1.04 \le m < 1.2$, or deep out of the money (DOTM) if $m \ge 1.2$. An option contract can be classified by time to maturity as short maturity (<60 days), medium maturity (60–160 days), or long maturity (>160 days). Table 1 describes the 101,036 call option prices and the implied volatilities used in the empirical analysis. The average call price ranges from $354.06 for long-maturity DITM options to $0.28 for short-maturity DOTM options. ITM, ATM, and OTM options account for 20, 22, and 21 percent of the total sample, respectively. Short- and long-maturity options account for 37 and 20 percent of the total sample, respectively. The table also shows the volatility smile and the corresponding term structure. For each set of maturities, the smile across moneyness is evident. The longer the time to maturity, the flatter the volatility smile.

5.2 Implementing the Option Pricing Models

All nonparametric regressions use the local linear approximation and the Epanechnikov kernel. We set the bandwidth at $h = 0.3s$, where s is the sample standard deviation of the moneyness. We also experimented with other bandwidth values and found that the overall pricing results for all nonparametric and semiparametric models were largely the same. In all time-weighted regressions, we set $\lambda = 0.83$ and measured the distance $|t - t_0|$ in calendar days. This choice assigns the weights 0.68, 0.83, 1, 0.83, and 0.68 to options traded each day

Table 2. In-sample pricing errors

	Bias	RMSE	MADE	Min	Max	Err>0%	Bias (%)	RMSE (%)	MADE (%)
Panel A: Aggregated valuation errors across all years									
ACE	0.02	0.38	0.27	−1.81	2.02	51.62	0.92	8.20	3.60
Semip-BS	0.07	0.48	0.30	−3.44	3.05	57.91	1.92	9.62	4.01
Ad Hoc BS	0.17	1.14	0.80	−4.05	5.31	51.08	9.24	38.54	17.20
NP	0.62	1.19	0.82	−2.57	5.61	68.99	5.51	18.66	9.63
GARCH	−0.44	1.25	0.90	−7.99	6.27	27.14	−18.34	42.78	21.70
Panel B: Valuation errors by years									
2002									
ACE	0.06	0.37	0.25	−1.81	2.02	54.20	1.02	9.80	4.78
Semip-BS	0.11	0.61	0.37	−3.44	3.05	58.22	2.13	11.42	5.55
Ad Hoc BS	0.10	1.15	0.77	−4.05	5.31	50.20	7.83	40.64	21.38
NP	0.83	1.35	0.89	−1.34	5.61	73.91	6.39	20.86	12.04
GARCH	−0.52	1.28	0.91	−7.99	4.36	22.36	−29.60	51.84	31.69
2003									
ACE	0.01	0.34	0.25	−1.69	1.43	51.44	0.91	8.27	3.64
Semip-BS	−0.02	0.36	0.24	−1.74	1.17	52.39	1.37	8.20	3.32
Ad Hoc BS	0.07	0.98	0.71	−3.30	4.08	48.96	7.22	30.91	14.55
NP	0.63	1.07	0.74	−1.04	4.27	72.18	4.98	16.02	8.60
GARCH	−0.39	0.93	0.72	−4.62	4.92	26.56	−20.21	42.06	21.36
2004									
ACE	0.01	0.42	0.31	−1.61	1.91	49.30	0.83	6.19	2.43
Semip-BS	0.12	0.43	0.28	−1.38	2.05	62.81	2.24	8.95	3.19
Ad Hoc BS	0.33	1.27	0.90	−3.15	5.02	53.92	12.50	42.73	15.70
NP	0.42	1.13	0.83	−2.57	3.93	61.26	5.18	18.75	8.29
GARCH	−0.40	1.45	1.06	−5.89	6.27	32.27	−5.78	32.65	12.45

NOTE: Shown are the bias, RMSE, and mean absolute error (MADE) of the dollar pricing error (model price − market price), and of the percentage relative pricing error, 100 × (model price − market price)/market price; Min (Max) is the minimum (maximum) dollar pricing error. Err>0% is the percentage of positive pricing errors for the different pricing models and call option prices from January 2, 2002 to December 31, 2004.

on Monday through Friday, respectively. We also experimented using other values for λ and obtained very similar results for all of the pricing models. In practice, the choice of λ is an empirical issue, but the more options that are available on day t_0, the lower λ should be. As in the GLR test (see Sec. 3), we estimated the correction term $\bar{F}_{t,c}$ in the ACE method (11) from the 0.05 to the 0.95 quantiles of the observed moneyness, and beyond this interval, we set $\bar{F}_{t,c}$ to 0.

When implementing the GARCH pricing formula (22), the dividends paid by the stocks in the S&P 500 index must be taken into account. For each maturity, we computed the dividend yield using spot-forward and put-call parities based on liquid closest ATM options, then subtract this yield from the current index level. All other pricing methods are implemented using forward prices, in which dividend yields are already embedded.

5.3 In-Sample Model Comparisons

For each Wednesday between January 2, 2002 and December 31, 2004, we calibrated the GARCH pricing model (21) to the cross-section of options. Aggregating data over each week, each Wednesday and for each maturity, we fit our ACE approach (11), the direct nonparametric model (7), the semiparametric Black–Scholes model (20), and the ad hoc Black–Scholes model (8). We then priced the European options available on each Wednesday, obtaining 16,521 option price estimates for each model.

Table 2 summarizes the pricing errors of the five option pricing models across the different years from 2002 to 2004. Overall, the ACE method has the best pricing performance; for example, it has a 67% lower RMSE than the benchmark ad hoc Black–Scholes model. Tables 3 and 4 disaggregate dollar and relative pricing errors across the five moneyness and three maturities categories. The ACE method has the lowest RMSE in most comparisons. It has some difficulties in pricing DITM options, because to price options with low moneyness (e.g., $m \approx 0.4$), the survivor function must be integrated for almost all of the moneyness domain, accumulating the estimation errors in $\bar{F}_t(m)$. Interestingly, the ACE method performs well in pricing ATM and OTM options, which are more actively traded than ITM options. These results demonstrate that the two key characteristics of our method are very effective in producing accurate pricing results. On the one hand, they confirm that the spaces of state price distributions are better suited for pricing options than the spaces of state price densities. This is because using state price distributions avoids a numerical differentiation (performing instead an integration by parts), which induces a more stable numerical procedure. On the other hand, our results show that the model-guided nonparametric procedure allows one to recover state price distributions precisely and thus pricing options accurately.

As shown in Tables 2, 3, and 4, the second-best method is the semiparametric Black–Scholes model (20) (label Semip-BS), which performs particularly well for DITM options. Although the ad hoc Black–Scholes model (8) (label Ad Hoc BS)

Table 3. In-sample dollar pricing errors disaggregated by moneyness and maturity

		Maturity					
		Less than 60		60 to 160		More than 160	
		Bias	RMSE	Bias	RMSE	Bias	RMSE
DITM	ACE	−0.04	0.49	−0.03	0.47	0.06	0.55
	Semip-BS	−0.01	0.08	0.00	0.16	0.01	0.25
	Ad Hoc BS	−0.06	0.17	−0.18	0.43	−0.33	0.60
	NP	−0.24	0.59	−0.22	0.68	0.02	0.55
	GARCH	−0.39	0.84	−0.72	1.37	−1.08	2.63
ITM	ACE	−0.08	0.44	−0.13	0.42	−0.10	0.42
	Semip-BS	−0.00	0.33	0.04	0.48	0.06	0.63
	Ad Hoc BS	−0.54	0.78	−0.99	1.26	−0.94	1.26
	NP	0.15	0.72	0.32	0.91	0.59	1.01
	GARCH	−0.63	1.16	−1.41	1.66	−2.05	2.40
ATM	ACE	0.01	0.39	−0.01	0.38	0.01	0.37
	Semip-BS	0.11	0.62	0.16	0.74	0.16	0.82
	Ad Hoc BS	0.80	1.56	0.36	1.43	0.20	1.33
	NP	1.43	1.75	1.55	1.90	1.32	1.69
	GARCH	0.75	1.40	−0.11	0.73	−0.18	1.05
OTM	ACE	0.14	0.28	0.24	0.37	0.28	0.43
	Semip-BS	0.11	0.39	0.14	0.50	0.19	0.66
	Ad Hoc BS	0.89	1.29	1.17	1.51	0.92	1.44
	NP	0.73	1.11	1.07	1.40	1.21	1.54
	GARCH	−0.18	0.82	−0.60	0.89	0.01	0.93
DOTM	ACE	−0.02	0.06	0.01	0.09	0.06	0.14
	Semip-BS	0.00	0.07	0.02	0.11	0.01	0.20
	Ad Hoc BS	−0.06	0.19	0.12	0.37	0.16	0.35
	NP	−0.03	0.10	0.07	0.25	0.16	0.34
	GARCH	−0.29	0.32	−0.41	0.53	−0.52	0.65

NOTE: Shown are the bias and RMSE of the pricing error (model price − market price), for the different models and call option prices from January 2, 2002 to December 31, 2004. See Table 1 for the definitions of DITM, ITM, ATM, OTM, and DOTM. Maturity is measured in calendar days.

is estimated for each maturity, it does not perform well, mainly because volatility smiles are not well approximated by parabolic functions. The semiparametric Black–Scholes model is designed to circumvent this problem, and it always outperforms the ad hoc Black–Scholes model. Table 5 gives summary statistics for the parabola coefficients, confirming that the volatility smile is a stable and persistent characteristic of implied volatilities.

The direct nonparametric model (7) (label NP) does not perform well, as anticipated, because it does not exploit the prior knowledge on the shape of the state price survivor function, and because a constant bandwidth is not adequate for estimating the survivor function over the whole domain. Comparing the ACE and NP methods shows that the proposed bias reduction technique is very effective in reducing pricing errors. Tables 2, 3, and 4 provide stark evidence of the power of our idea of combining model-based pricing formulas and nonparametric learning to correct pricing errors.

Compared with the other methods, the GARCH pricing model (21) (label GARCH) is not sufficiently flexible to fit options with different maturities. Table 6 gives the calibrated GARCH parameters. As reported previously (e.g., Heston and Nandi 2000; Barone-Adesi, Engle, and Mancini 2008), the parameter γ is largely positive, confirming that negative shocks increase the volatility more than positive shocks of same absolute magnitude. The GARCH parameters change over time, but the long-run volatility and the persistency of the variance process, $\beta + \alpha\gamma^2$, remain quite stable.

Figure 5 shows the absolute dollar and relative pricing errors across moneyness and maturity categories for the different pricing models. The two proposed methods—the ACE and the semiparametric Black–Scholes methods—outperform all of the other pricing models. In most situations, the ACE method outperforms the semiparametric Black–Scholes method. Given that the direct nonparametric model is largely outperformed by the ACE method, its graph is omitted.

Nonparametric estimations of state price survivor functions in the ACE method are guided by the parametric model (9), which addresses the issue of the volatility smile. An interesting question is how the ACE method behaves when a simpler, less accurate model is used as a parametric start. A naive parametric start is zero, resulting in the direct nonparametric method. This method does not work well, as shown in Figure 3 and Tables 2–4. A more interesting question, raised by a referee, is the behavior of the ACE with the Black–Scholes model as a parametric start. To address this question, we repeated the previous in-sample (and the subsequent out-of-sample and hedging) analysis, replacing model (9) with the lognormal survivor function as in the Black–Scholes model. To save space, we omit these additional results here, but they are available on request. As expected, the overall performance of this modified ACE method is less accurate, because the Black–Scholes model is

Table 4. In-sample percentage relative pricing errors disaggregated by moneyness and maturity

		Maturity					
		Less than 60		60 to 160		More than 160	
		Bias	RMSE	Bias	RMSE	Bias	RMSE
DITM	ACE	−0.01	0.18	−0.01	0.16	0.01	0.18
	Semip-BS	0.00	0.04	0.00	0.07	0.00	0.10
	Ad Hoc BS	−0.03	0.08	−0.08	0.18	−0.13	0.24
	NP	−0.07	0.20	−0.06	0.23	0.02	0.20
	GARCH	−0.15	0.34	−0.28	0.51	−0.43	0.89
ITM	ACE	−0.08	0.39	−0.12	0.33	−0.08	0.29
	Semip-BS	0.00	0.35	0.04	0.43	0.04	0.49
	Ad Hoc BS	−0.50	0.79	−0.77	1.04	−0.65	0.92
	NP	0.20	0.72	0.30	0.78	0.44	0.78
	GARCH	−0.49	1.02	−1.03	1.20	−1.33	1.55
ATM	ACE	0.78	2.69	0.21	1.25	0.09	0.76
	Semip-BS	1.02	4.12	0.56	2.25	0.32	1.41
	Ad Hoc BS	7.72	16.38	2.16	5.77	0.73	2.69
	NP	8.60	13.26	4.47	6.38	2.37	3.25
	GARCH	4.32	13.07	0.00	2.18	−0.12	1.62
OTM	ACE	5.77	14.48	4.95	9.54	2.48	4.49
	Semip-BS	8.50	20.24	4.54	12.03	1.98	6.40
	Ad Hoc BS	54.33	88.42	34.56	57.92	10.95	21.00
	NP	26.02	43.59	18.18	26.86	8.96	12.51
	GARCH	−31.46	66.62	−24.90	38.96	−4.38	17.05
DOTM	ACE	−6.20	18.66	−0.74	16.20	1.84	11.62
	Semip-BS	1.70	17.77	4.54	16.72	2.39	12.84
	Ad Hoc BS	−25.84	60.48	6.38	57.80	13.27	40.12
	NP	−13.58	28.18	−0.02	24.55	4.84	17.39
	GARCH	−97.78	100.47	−91.64	93.28	−66.10	74.60

NOTE: Shown are the bias and RMSE of the relative pricing error, $100 \times (\text{model price} - \text{market price})/\text{market price}$, for the different models and call option prices from January 2, 2002 to December 31, 2004. See Table 1 for the definitions of DITM, ITM, ATM, OTM, and DOTM. Maturity is measured in calendar days.

less accurate. Interestingly, this ACE method still outperforms the ad hoc Black–Scholes model and the direct nonparametric approach, again demonstrating that the nonparametric corrections are highly important and outperform the parametric method used initially. This also indicates that a better parametric model puts less burden on the nonparametric correction and thus yields more accurate pricing. In particular, if the initial pricing formula is zero, namely the worst parametric model, then the resulting method is direct nonparametric estimation. Thus the direct method is not expected to perform well. Our empirical comparisons confirm the advantages of using model-guided nonparametric estimations over direct nonparametric estimation.

In previous analyses, the nonparametric and semiparametric methods were estimated on each Wednesday using also future information, namely options data on Thursday and Friday. This approach exploits the time continuity of the pricing function (as in, e.g., Aït-Sahalia and Lo 1998); however, these methods also can be implemented using only current and past data. As a robustness check, we repeated the in-sample (and the subsequent out-of-sample and hedging) analysis pricing options on each Friday using data from Monday to Friday without borrowing future information. These additional results (not reported here, but available on request) largely confirm the findings reported here. As a further robustness check, we also repeated the in-sample (and the subsequent out-of-sample and hedging) analysis of the ACE method using an alternative kernel, namely the triweight kernel. The results were nearly the same as those based on the Epanechnikov kernel and are omitted, but are available on request.

Table 5. Ad hoc Black–Scholes model (8), with mean and SD of model parameters calibrated each Wednesday from January 2, 2002 to December 31, 2004, using call option prices

	a_0		a_1		a_2	
Year	Mean	SD	Mean	SD	Mean	SD
2002	0.84	0.09	−0.99	0.15	0.38	0.07
2003	0.76	0.08	−0.90	0.18	0.34	0.10
2004	0.81	0.10	−1.06	0.21	0.41	0.11

5.4 Out-of-Sample Model Comparisons

Out-of-sample pricing of options is an interesting challenge for any pricing method. It not only tests the goodness-of-fit of the pricing formula, but also checks whether or not the method overfits the option prices in the in-sample estimation. On each Wednesday and for each maturity, we used in-sample model estimates to price the same options 1 week later using forward prices, times to maturity, and interest rates relevant on the next Wednesday. Dumas, Fleming, and Whaley (1998) and Heston

Table 6. Heston and Nandi GARCH model (21), with the mean and SD of model parameters (daily base) calibrated each Wednesday from January 2, 2002 to December 31, 2004, using call option prices

	$\omega \times 10^{14}$		β		γ		$\alpha \times 10^{6}$		$\sqrt{E[h]}$		$\beta + \alpha\gamma^2$	
Year	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
2002	2.53	7.60	0.67	0.14	246.15	73.27	5.79	3.66	0.23	0.04	0.96	0.03
2003	1.07	3.96	0.66	0.22	284.69	93.57	5.48	6.89	0.22	0.03	0.95	0.07
2004	0.82	3.09	0.64	0.14	563.63	361.61	2.41	2.71	0.19	0.05	0.96	0.06

NOTE: Risk-neutral long-run volatility (annual base) $\sqrt{E[h]} = \sqrt{365(\omega + \alpha)/(1 - \beta - \alpha\gamma^2)}$, and persistence of the variance process $\beta + \alpha\gamma^2$.

and Nandi (2000), among others, also used a 1-week-ahead forecast horizon. In the ACE method, the model-based pricing formula, $\bar{F}_{\mathrm{LN}}(m; \vartheta_t)$, accounts for the shorter time to maturity. For this purpose, we fit the following regression model:

$$\hat{\vartheta}_t = \kappa\sqrt{T - t} + \mathrm{error}_t \quad \text{for } t \in [t_0 - d, t_0 + d], \tag{23}$$

where $\hat{\vartheta}_t$s are as given in (10). Equation (23) links times to maturity and $\hat{\vartheta}_t$s in the state price survivor function, $\bar{F}_{\mathrm{LN}}(m; \hat{\vartheta}_t)$. Using in-sample data observed over 1 trading week, the least-squares estimate gives

$$\hat{\kappa} = \frac{\sum_{i=1}^{5} \hat{\vartheta}_i \sqrt{\tau_i}}{\sum_{i=1}^{5} \tau_i}.$$

The coefficient $\hat{\vartheta}_{t_1} = \hat{\kappa}\sqrt{T - t_1}$ reflects the shorter time to maturity at the future date t_1. In our application, $t_1 = t_0 + 7$ days. The $\hat{\vartheta}_{t_1}$ could be easily estimated by solving the minimization problem (10) using options data on time t_1, but this procedure would invalidate the out-of-sample analysis.

The out-of-sample pricing performances are summarized in Table 7 and disaggregated by moneyness and maturity in Tables 8 and 9; see also Figure 6. Interestingly, these results exhibit the same pattern as in the in-sample analysis. The ACE method outperforms all of the other pricing methods in terms of dollar and relative pricing errors, sometimes by an even greater extent. These results demonstrate that the automatic bias correction in the ACE is very effective. Without this feature, the ACE method is essentially the same as the ad hoc Black–Scholes method, which does not perform well. These results also demonstrate that the second-best method is our semiparametric Black–Scholes model, although the relative pricing errors increase when times to maturity decrease. In the GARCH model, the conditional variance, h_t, is updated using the risk-neutral parameters calibrated at t_0 and the actual S&P 500 daily log-returns from t_0 to $t_0 + 7$ days. Tables 7, 8, and 9 show that the GARCH model has larger prediction errors than the ACE and the semiparametric Black–Scholes methods. These findings confirm the power of nonparametric approaches. We omit the direct nonparametric model, because it was largely outperformed by the ACE method in the in-sample analysis.

As a robustness check, we repeated the previous out-of-sample analysis using a 1-month horizon. The results exhibit the same pattern as those reported in Tables 7, 8, and 9 and thus are omitted here, but are available on request. In this out-of-sample analysis, the GARCH model becomes slightly more competitive and tends to outperform the ad hoc Black–Scholes model more often, but is still outperformed by our ACE method.

5.5 Hedging Results

An important motivation for developing option pricing models is to provide better risk management of the derivatives assets. Setting hedge ratios based on an accurate and reliable valuation model should result in improved hedging performance. Following Dumas, Fleming, and Whaley (1998), we evaluated the performance of a hedge portfolio formed on day t and liquidated 1 week later on day $t + 7$. The return on such a discretely adjusted hedge portfolio has three components: the risk-free return on investment, the return from the discrete adjustment of the hedge, and the return from the difference between the change in the actual option price and the change in the theoretical option price over the 1-week horizon (see Galai 1983; Dumas, Fleming, and Whaley 1998). We use forward option prices, and thus the risk-free return component of the hedge portfolio is 0. Because the focus is on model performance, not on the issues raised by discrete time rebalancing, we assume that the hedge portfolio is rebalanced continuously over time (see, e.g., Bossaerts and Hillion 1997, 2003 for hedging analyses in discrete time). Thus the hedging error is defined as

$$\epsilon_t = \Delta C_{\mathrm{actual},t} - \Delta C_{\mathrm{model},t}, \tag{24}$$

where $\Delta C_{\mathrm{actual},t}$ is the observed change in the market option price from date t to date $t+7$ and $\Delta C_{\mathrm{model},t}$ is the change in the model theoretical price. The proof of Equation (24) has been provided by Dumas, Fleming, and Whaley (1998, sec. VI), but for completeness we recall it here as well. The hedging error resulting from the continuous time recalibration using the hedge ratio, ρ, is

$$\Delta C_{\mathrm{actual},t} - \int_t^{t+7} \rho(S_u, u)\, dS_u. \tag{25}$$

If the valuation model gives the correct hedge ratio, ρ, to continuously rebalance the hedge, then the two terms in (25) will be equal with probability 1, and the hedging error, ϵ_t, will be 0.

Table 10 summarizes the hedging errors of the four pricing models, and Table 11 disaggregates the hedging results across moneyness and maturity. These hedging results largely confirm the previous in-sample and out-of-sample pricing results. Our ACE approach tends to outperform all the other methods, sometimes significantly. The overall second-best method is our semiparametric Black–Scholes model. The ad hoc Black–Scholes model tends to dominate the GARCH model. Figure 7 shows the absolute hedging errors of the pricing models and visually confirms the previous results.

Another interesting exercise is to use the different models to hedge calendar spread portfolios. The rationale for this hedging

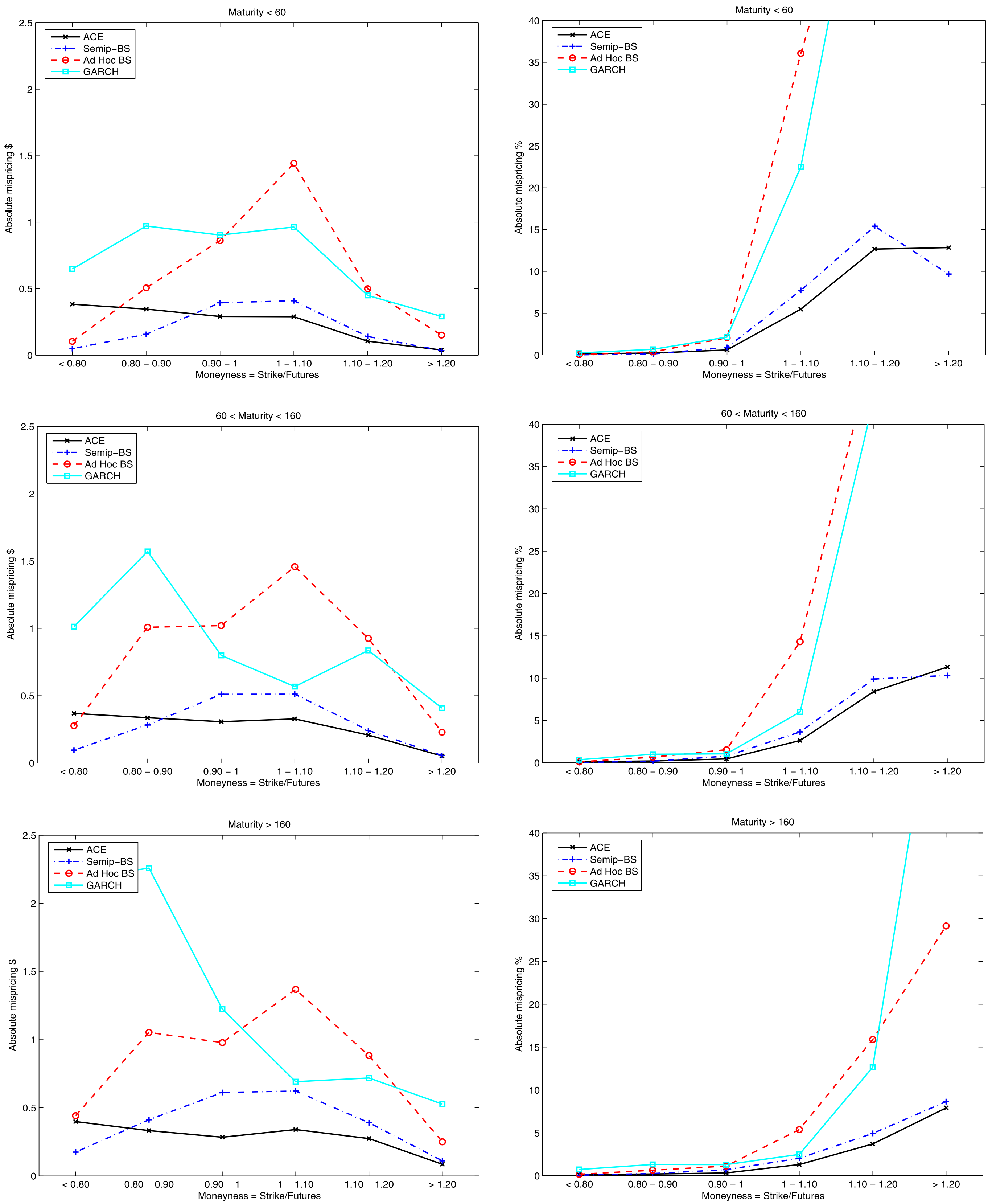

Figure 5. In-sample absolute mispricing in dollars, that is, |model price − market price| (left graphs), and in percentage, i.e., 100 × |model price − market price|/market price (right graphs), for the different pricing models, averaged across the Wednesdays from January 2, 2002 to December 31, 2004 for the SPX call options.

Table 7. Out-of-sample pricing errors

	Bias	RMSE	MADE	Min	Max	Err>0%	Bias (%)	RMSE (%)	MADE (%)
Panel A: Aggregated valuation errors across all years									
ACE	0.01	1.01	0.64	−5.31	9.81	51.29	0.87	24.96	11.44
Semip-BS	0.03	1.48	0.90	−7.21	6.89	44.66	−2.06	25.15	12.92
Ad Hoc BS	0.11	1.77	1.15	−7.86	7.17	45.62	6.69	47.66	22.05
GARCH	−0.49	2.27	1.44	−16.97	9.00	30.91	−18.69	48.20	25.66
Panel B: Valuation errors by years									
2002									
ACE	−0.01	1.24	0.77	−5.31	9.36	48.93	0.17	31.64	16.95
Semip-BS	−0.22	1.93	1.21	−7.21	6.89	36.30	−7.90	32.68	19.97
Ad Hoc BS	−0.23	2.12	1.38	−7.86	6.97	38.73	−1.25	49.25	27.53
GARCH	−0.86	2.62	1.65	−14.58	8.63	22.46	−32.92	55.23	37.09
2003									
ACE	−0.06	0.88	0.54	−3.92	9.81	46.03	−1.69	21.63	9.64
Semip-BS	0.05	1.13	0.71	−5.55	4.14	47.52	−2.15	20.58	10.57
Ad Hoc BS	0.12	1.43	0.94	−6.21	5.09	46.42	4.77	36.87	18.00
GARCH	−0.39	1.74	1.16	−9.39	9.00	28.90	−21.01	43.69	23.57
2004									
ACE	0.10	0.90	0.61	−4.79	3.12	58.47	3.95	20.78	8.23
Semip-BS	0.24	1.30	0.81	−5.51	5.80	49.45	3.29	21.02	8.82
Ad Hoc BS	0.41	1.73	1.16	−4.99	7.17	51.04	15.66	54.84	21.00
GARCH	−0.27	2.38	1.53	−16.97	6.51	40.45	−3.66	45.45	17.39

NOTE: Shown are the bias, RMSE, and MADE of the dollar pricing error (model price − market price), and of the percentage relative pricing error, 100 × (model price − market price)/market price; Min (Max) is the minimum (maximum) dollar pricing error. Err>0% is the percentage of positive pricing errors for the different models and call option prices from January 2, 2002 to December 31, 2004.

analysis is that on each day t, the ACE, semiparametric, and ad hoc Black–Scholes models are separately fitted to each time to maturity, whereas the GARCH model is calibrated to the entire cross-section of options, automatically imposing a time consistency across the estimated state price survivor functions. For each day t, we consider a calendar spread portfolio consisting

Table 8. Out-of-sample dollar pricing errors disaggregated by moneyness and maturity

		Maturity					
		Less than 60		60 to 160		More than 160	
		Bias	RMSE	Bias	RMSE	Bias	RMSE
DITM	ACE	0.08	0.40	−0.08	0.52	−0.18	0.67
	Semip-BS	−0.08	0.25	−0.08	0.55	−0.02	0.88
	Ad Hoc BS	−0.11	0.30	−0.27	0.72	−0.37	1.02
	GARCH	−0.28	0.75	−0.85	2.04	−1.37	4.35
ITM	ACE	−0.18	0.71	−0.43	1.19	−0.40	1.40
	Semip-BS	−0.10	1.01	−0.08	1.58	−0.04	2.18
	Ad Hoc BS	−0.55	1.21	−1.10	1.94	−0.98	2.36
	GARCH	−0.55	1.22	−1.58	2.91	−2.39	5.40
ATM	ACE	0.05	1.17	0.01	1.58	−0.13	1.67
	Semip-BS	0.33	1.79	0.35	2.37	0.17	2.89
	Ad Hoc BS	0.87	2.26	0.44	2.59	0.20	3.08
	GARCH	0.76	1.87	−0.19	2.59	−0.51	4.58
OTM	ACE	0.27	0.74	0.44	1.30	0.48	1.25
	Semip-BS	−0.04	1.09	0.06	1.59	0.11	2.11
	Ad Hoc BS	0.58	1.49	1.04	2.09	0.83	2.60
	GARCH	−0.24	1.06	−0.76	1.74	−0.34	2.66
DOTM	ACE	−0.08	0.15	−0.01	0.27	0.11	0.38
	Semip-BS	−0.13	0.19	−0.09	0.32	−0.17	0.62
	Ad Hoc BS	−0.15	0.23	0.02	0.39	−0.03	0.66
	GARCH	−0.27	0.31	−0.43	0.62	−0.65	1.02

NOTE: Shown are the bias and RMSE of the pricing error (model price − market price), for the different models and call option prices from January 2, 2002 to December 31, 2004. See Table 1 for the definitions of DITM, ITM, ATM, OTM, and DOTM. Maturity is measured in calendar days.

Table 9. Out-of-sample percentage relative pricing errors disaggregated by moneyness and maturity

		Maturity					
		Less than 60		60 to 160		More than 160	
		Bias	RMSE	Bias	RMSE	Bias	RMSE
DITM	ACE	0.02	0.16	−0.04	0.22	−0.07	0.27
	Semip-BS	−0.03	0.12	−0.03	0.25	−0.01	0.36
	Ad Hoc BS	−0.04	0.14	−0.12	0.32	−0.15	0.42
	GARCH	−0.11	0.31	−0.34	0.78	−0.56	1.58
ITM	ACE	−0.19	0.74	−0.34	1.07	−0.27	1.03
	Semip-BS	−0.09	1.05	−0.05	1.38	−0.03	1.63
	Ad Hoc BS	−0.50	1.25	−0.85	1.64	−0.67	1.75
	GARCH	−0.45	1.15	−1.15	2.25	−1.54	3.71
ATM	ACE	2.92	11.63	0.65	4.56	−0.03	2.82
	Semip-BS	4.01	14.64	1.54	6.77	0.53	4.91
	Ad Hoc BS	11.12	27.70	2.82	8.82	0.89	5.44
	GARCH	6.55	22.42	0.19	6.63	−0.39	7.17
OTM	ACE	8.53	45.11	11.32	27.59	5.48	11.31
	Semip-BS	−1.19	43.20	6.58	27.59	3.52	15.59
	Ad Hoc BS	46.17	99.17	36.30	71.46	12.35	28.19
	GARCH	−28.49	85.41	−24.21	41.23	−5.51	21.53
DOTM	ACE	−23.22	59.06	−6.63	48.31	4.87	24.52
	Semip-BS	−41.29	60.32	−14.06	42.23	−7.65	28.81
	Ad Hoc BS	−52.94	77.19	−8.61	67.63	2.91	49.15
	GARCH	−97.94	101.48	−92.26	93.98	−68.63	76.91

NOTE: Shown are the bias and RMSE of the relative pricing error, $100 \times (\text{model price} - \text{market price})/\text{market price}$, for the different models and call option prices from January 2, 2002 to December 31, 2004. See Table 1 for the definitions of DITM, ITM, ATM, OTM, and DOTM. Maturity is measured in calendar days.

of a long position in a call option with the longest time to maturity available in our database and a short position in a call option with same strike price and shortest available time to maturity. We assume that the portfolio can be continuously hedged over time and is liquidated 1 week later at time $t + 7$. We compute hedging errors for all calendar spread portfolios available in our database and for each option pricing model. Overall, these findings confirm the previous hedging results and are not reported here, but are available on request. The hedging performance of the GARCH model seems to improve somehow compared with previous hedging results, but the flexibility of separately fitting options with different maturities outweighs the time consistency of the GARCH model.

6. CONCLUSIONS

We have proposed a new nonparametric method for estimating state price distributions and pricing financial derivatives. This method, called the automatic correction of errors (ACE) in a pricing formula, is based on a model-guided nonparametric estimate of the state price survivor function. Given any parametric pricing model, the induced pricing errors are nonparametrically learned and corrected thorough the estimate of the survivor function, improving the model pricing performance. Our ACE method is easy to implement and can be combined with any model-based pricing formula to correct the systematic biases of pricing errors. We also have proposed a semiparametric Black–Scholes method for option pricing, simplifying the method introduced by Aït-Sahalia and Lo (1998). Empirical studies based on S&P 500 index options show that the ACE approach outperforms, in terms of predictive and hedging abilities, the ad hoc Black–Scholes, the semiparametric Black–Scholes, the direct nonparametric, and the GARCH option pricing models. Our proposed GLR test demonstrates that the ACE method is very effective in reducing pricing errors. Our ACE approach could be applied in other contexts as well. For example, in credit risk modeling, accurate estimation of the default survivor function is essential in the pricing of credit risk-sensitive contingent claims, and this accuracy could be achieved by applying the ACE method.

APPENDIX A: PROOF OF PROPOSITION 1

Let $\bar{m} = (m_1 + m_2)/2$. By (2), omitting subscripts t, the left side of (5) can be expressed as

$$\frac{1}{m_2 - m_1}\int_{m_1}^{m_2} \bar{F}(u)\,du = \bar{F}(\bar{m}) + \frac{1}{m_2 - m_1}\int_{m_1}^{m_2} (\bar{F}(u) - \bar{F}(\bar{m}))\,du.$$

We now evaluate the approximation error. Because $\bar{F} = 1 - F$, $\bar{F}' = -f$, and $\bar{F}'' = -f'$, by a Taylor expansion to the second order, the second integral can be expressed as

$$\int_{m_1}^{m_2}\left(-f(\bar{m})(u-\bar{m}) - \frac{1}{2}f'(\xi)(u-\bar{m})^2\right)du,$$

where ξ is a point lying between m_1 and m_2. The first term is 0, which is an advantage of using the midpoint, $\bar{m}$, and the second integral is bounded by

$$\frac{1}{2}\left\{\max_{m_1\le\xi\le m_2} -f'(\xi)\right\}\int_{m_1}^{m_2}(u-\bar{m})^2\,du$$
$$= -\frac{1}{24}\left\{\min_{m_1\le\xi\le m_2} f'(\xi)\right\}(m_2 - m_1)^3.$$

This completes the derivation of Proposition 1.

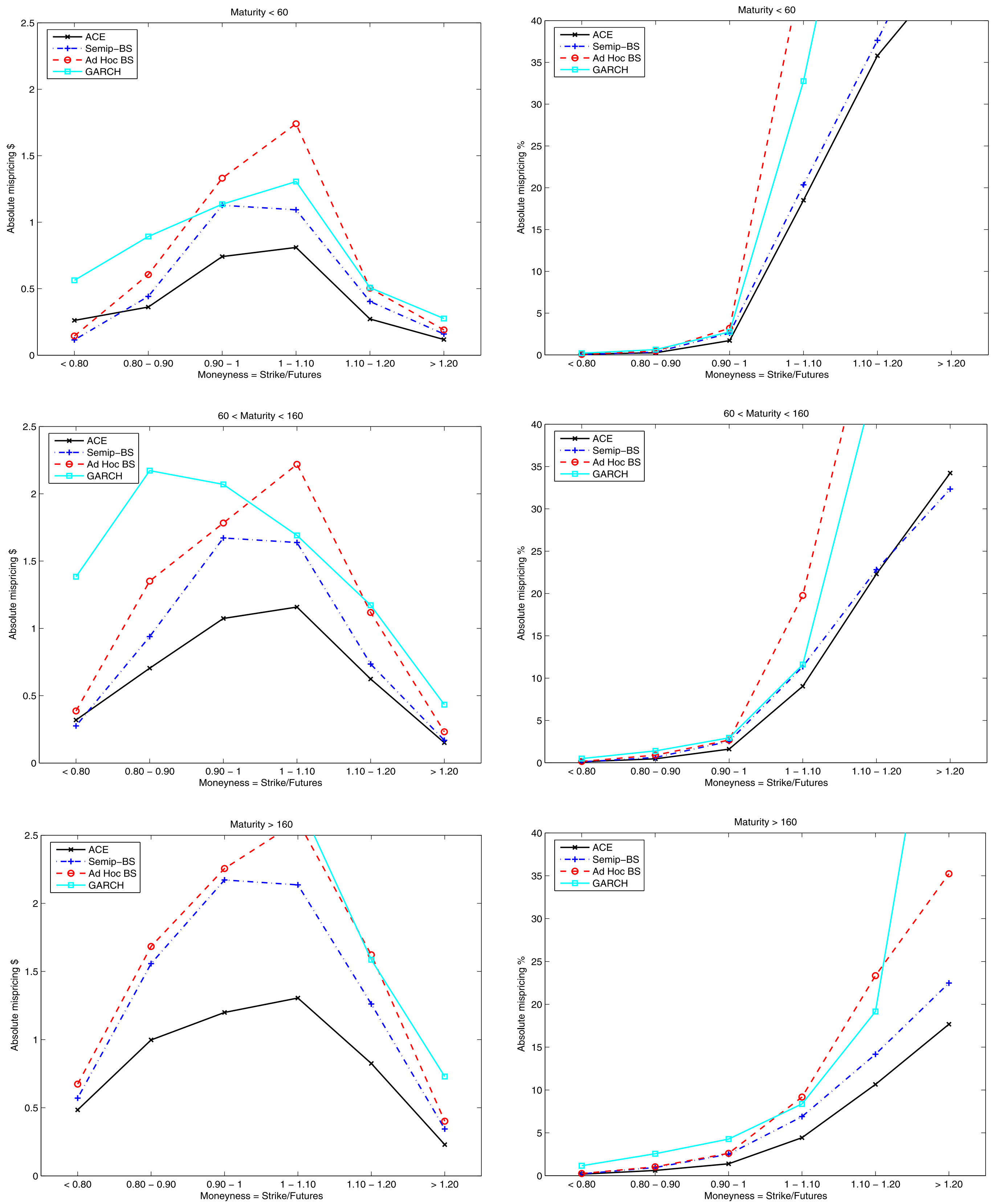

Figure 6. Out-of-sample absolute mispricing in dollars, that is, |model price − market price| (left graphs), and in percentage, i.e., 100 × |model price − market price|/market price (right graphs), for the different pricing models, averaged across the Wednesdays from January 2, 2002 to December 31, 2004 for the SPX call options.

Table 10. Hedging error

	Mean	RMSE	MADE	Min	Max	Err>0%
Panel A: Aggregated hedging errors across all years						
ACE	−0.00	0.47	0.32	−3.01	2.73	51.61
Semip-BS	0.01	0.62	0.40	−4.07	3.00	53.35
Ad Hoc BS	−0.02	0.67	0.45	−4.11	3.15	51.92
GARCH	−0.06	1.15	0.73	−10.87	9.56	46.13
Panel B: Hedging errors by years						
2002						
ACE	0.01	0.49	0.31	−3.01	2.47	52.60
Semip-BS	0.03	0.77	0.49	−4.07	3.00	56.14
Ad Hoc BS	−0.02	0.80	0.54	−4.11	2.75	55.82
GARCH	−0.07	1.22	0.76	−10.87	9.56	45.26
2003						
ACE	−0.02	0.41	0.29	−1.87	1.87	51.21
Semip-BS	−0.01	0.51	0.33	−2.13	2.03	50.40
Ad Hoc BS	−0.04	0.57	0.40	−2.63	2.01	49.46
GARCH	−0.06	0.89	0.56	−7.00	5.28	46.42
2004						
ACE	0.00	0.52	0.37	−1.92	2.73	51.02
Semip-BS	0.00	0.56	0.36	−2.42	2.94	53.43
Ad Hoc BS	−0.00	0.62	0.42	−2.19	3.15	50.42
GARCH	−0.05	1.28	0.86	−5.96	8.22	46.73

NOTE: Shown are the mean, RMSE, and MADE of the dollar hedging error in Equation (24). Min (max) is the minimum dollar hedging error. Err>0% is the percentage of positive hedging errors for the different models and call option prices from January 2, 2002 to December 31, 2004.

Table 11. Hedging errors disaggregated by moneyness and maturity

		Maturity					
		Less than 60		60 to 160		More than 160	
		Mean	RMSE	Mean	RMSE	Mean	RMSE
DITM	ACE	−0.01	0.73	0.01	0.65	0.04	0.71
	Semip-BS	0.00	0.13	0.01	0.23	0.03	0.37
	Ad Hoc BS	−0.00	0.19	0.00	0.31	0.02	0.44
	GARCH	−0.13	0.97	−0.02	1.30	−0.09	3.20
ITM	ACE	−0.02	0.64	−0.02	0.55	0.02	0.54
	Semip-BS	0.04	0.50	0.03	0.64	0.05	0.84
	Ad Hoc BS	−0.05	0.59	−0.02	0.69	0.08	0.86
	GARCH	−0.20	1.29	−0.21	0.91	−0.17	1.48
ATM	ACE	−0.01	0.48	−0.05	0.46	0.02	0.43
	Semip-BS	0.00	0.82	0.03	0.91	0.03	1.08
	Ad Hoc BS	−0.11	0.87	−0.03	0.87	0.05	1.02
	GARCH	0.03	1.33	−0.10	0.94	0.06	0.80
OTM	ACE	0.02	0.27	0.01	0.34	0.01	0.38
	Semip-BS	−0.05	0.52	−0.03	0.59	0.01	0.81
	Ad Hoc BS	0.01	0.69	−0.04	0.72	−0.03	0.85
	GARCH	0.02	0.88	0.02	0.72	0.00	1.02
DOTM	ACE	−0.01	0.08	−0.01	0.10	0.01	0.16
	Semip-BS	−0.01	0.13	−0.01	0.14	−0.00	0.24
	Ad Hoc BS	−0.01	0.24	−0.00	0.30	−0.01	0.29
	GARCH	−0.04	0.18	−0.04	0.23	−0.03	0.40

NOTE: Shown are the mean and RMSE of the dollar hedging error in Equation (24) for the different models and call option prices from January 2, 2002 to December 31, 2004. See Table 1 for the definitions of DITM, ITM, ATM, OTM, and DOTM. Maturity is measured in calendar days.

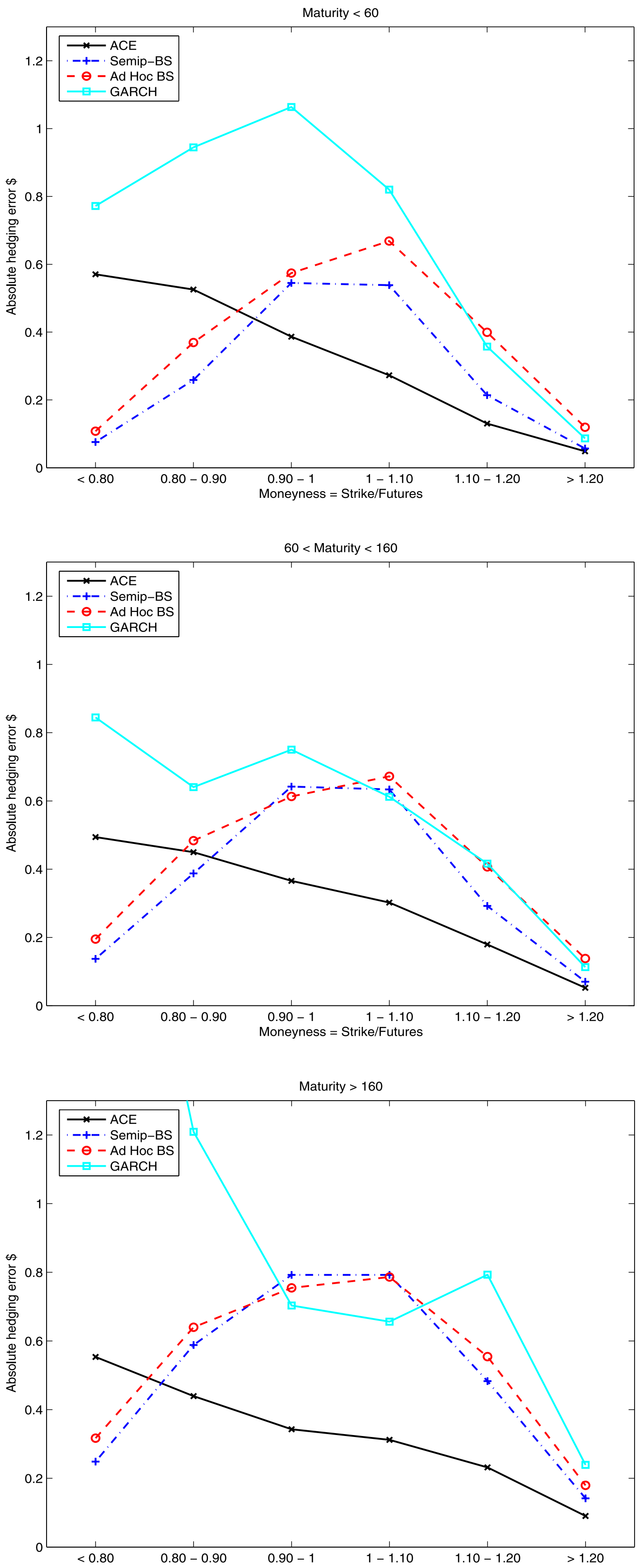

Figure 7. Absolute hedging error in dollars for the different pricing models, averaged across the Wednesdays from January 2, 2002 to December 31, 2004 for the SPX call options.

APPENDIX B: CONDITIONS AND PROOF OF PROPOSITION 3

To simplify the derivations, we make several idealizations. We assume that the data

$$\{(\bar{m}_{t,i}, Y_{t,i}), i = 1, \dots, N_t; t \in [t_0 - d, t_0 + d]\}$$

are a sequence of iid random variables satisfying model (6) with iid homoscedastic random noise, $\varepsilon_{t,i}$. In addition, we make the following technical assumptions:

- (B1) The marginal density $g(\cdot)$ of the moneyness $\{m_{t,i}\}$ is continuous at the point m. The conditional variance $\sigma_\varepsilon^2(\cdot)$ of $\varepsilon_{t,i}$ is continuous at the point m.
- (B2) $F_0(\cdot)$ has a continuous second derivative at the point m.
- (B3) $E|\varepsilon|^{2+\delta} < \infty$ for some $\delta > 0$.
- (B4) The function $K(t)$ is symmetric and has bounded support.
- (B5) The bandwidth h tends to 0 in such a way that $nh \to \infty$.
- (B6) The function $F(x; \theta)$ is Lipschitz-continuous in θ: $|F(x; \theta_1) - F(x; \theta_2)| \le C(x)|\theta_1 - \theta_2|$, with $C(x)$ bounded in a neighborhood of m. In addition, the calibrated $\hat{\theta}$ is root-n consistent.

From the definition of $\hat{\bar{F}}$ and $\bar{F}_c$, we have

$$\hat{\bar{F}}(m) - \bar{F}_0(m) = \bar{F}(m; \hat{\theta}) - \bar{F}(m; \theta_0) + \hat{\bar{F}}_c(m) - \bar{F}_c(m).$$

By condition (B6), the first difference term on the right side is of order $O_P(n^{-1/2})$. Thus

$$\hat{\bar{F}}(m) - \bar{F}_0(m) = \hat{\bar{F}}_c(m) - \bar{F}_c(m) + O_P(n^{-1/2}). \tag{B.1}$$

We now deal with the main term in (B.1). Write the local linear regression smoother as

$$\hat{\bar{F}}_c(m) = \sum_{t=t_0-d}^{t_0+d} \sum_{i=1}^{N_t} W_{t,i}(m) \tilde{Y}_{t,i}, \tag{B.2}$$

where $W_{t,i}(m)$ is the weight induced by the local linear regression (see, e.g., Fan and Yao 2003, section 6.3.3). The local linear weights satisfy (Fan and Yao 2003, section 6.3.3),

$$\sum_{t=t_0-d}^{t_0+d} \sum_{i=1}^{N_t} W_{t,i}(m) = 1, \qquad \sum_{t=t_0-d}^{t_0+d} \sum_{i=1}^{N_t} |W_{t,i}(m)| \le 2 + o(1). \tag{B.3}$$

Then, using (6), the data $\tilde{Y}_{t,i}$ can be written as

$$\tilde{Y}_{t,i} = \bar{F}_c(\bar{m}_{t,i}) + \varepsilon_{t,i} + Z_{t,i}, \tag{B.4}$$

where $Z_{t,i} = \bar{F}(\bar{m}_{t,i}; \theta_0) - \bar{F}(\bar{m}_{t,i}; \hat{\theta})$. By condition (B6), it follows that for a small neighborhood $\mathcal{N}$ around the point m,

$$\sup_{\bar{m}_{t,i} \in \mathcal{N}} |Z_{t,i}| = O_P(n^{-1/2}). \tag{B.5}$$

Because K has a bounded support, all data points that contribute to computing (B.2) fall in $\mathcal{N}$. Therefore, in (B.2), replacing $\tilde{Y}_{t,i}$ with (B.4) and using (B.3) and (B.5), we have

$$\hat{\bar{F}}_c(m) = \sum_{t=t_0-d}^{t_0+d} \sum_{i=1}^{N_t} W_{t,i}(m) \tilde{Z}_{t,i} + O_P(n^{-1/2}),$$

where $\tilde{Z}_{t,i} = \bar{F}_c(\bar{m}_{t,i}) + \varepsilon_{t,i}$. It follows from (B.1) that

$$\hat{\bar{F}}(m) - \bar{F}_0(m) = \hat{\bar{F}}_c^*(m) - \bar{F}_c(m) + O_P(n^{-1/2}),$$

where $\hat{\hat{F}}_c^*(m) = \sum_{t=t_0-d}^{t_0+d} \sum_{i=1}^{N_t} W_{t,i}(m)\tilde{Z}_{t,i}$ is the local linear regression smoother for the pseudodata

$$\{(\bar{m}_{t,i}, \tilde{Z}_{t,i}), i = 1, \ldots, N_t; t \in [t_0 - d, t_0 + d]\}.$$

The result follows from the asymptotic normality theory of the local linear regression (Fan and Yao 2003).

APPENDIX C: CONDITIONS AND PROOF OF PROPOSITION 4

Here we make the same idealized assumption as in Appendix B, that is,

$$\{(\bar{m}_{t,i}, Y_{t,i}), i = 1, \ldots, N_t; t \in [t_0 - d, t_0 + d]\}$$

is a sequence of iid random variables satisfying model (6) with iid homoscedastic random noise, $\varepsilon_{t,i}$. We further assume the following:

- (A1) The marginal density of moneyness $\{m_{t,i}\}$ is bounded away from 0 in the interval $[a, b]$.
- (A2) $F(\cdot)$ has a continuous second derivative.
- (A3) $E|\varepsilon|^4 < \infty$.
- (A4) The function $K(t)$ is symmetric and bounded. Furthermore, the functions $t^3K(t)$ and $t^3K'(t)$ are bounded, and $\int t^4 \times K(t)\,dt < \infty$.
- (A5) The bandwidth h satisfies $h \to 0$ and $nh^{3/2} \to \infty$.
- (A6) The function $F(x; \theta)$ has a continuous derivative with respect to θ, and the calibrated $\hat{\theta}$ is root-n consistent.

Similar assumptions have been made by Aït-Sahalia and Lo (1998) and Gagliardini, Gourieroux, and Renault (2005), among others.

Under the foregoing conditions, the number of observations in the interval $[a, b]$, $n_{a,b}$, by the central limit theorem, is

$$n^{-1}n_{a,b} = P(m \in [a, b]) + O_P(n^{-1/2}),$$

where n is the total number of sample observations (i.e., $n = \sum_{t=t_0-d}^{t_0+d} N_t$). Let $T_n' = \frac{n}{2}\log(\mathrm{RSS}_0/\mathrm{RSS}_1)$, we have

$$T_n = n^{-1}n_{a,b}T_n' = P(m \in [a, b])T_n' + O_P(n^{-1/2}T_n'). \tag{C.1}$$

Now we can apply the result of Fan, Zhang, and Zhang (2001) to T_n', noting that in our setting the null hypothesis is $\bar{F}_{t,c} = 0$. In particular, according to remark 4.2 of Fan, Zhang, and Zhang (2001), we have $r_K' T_n' \overset{a}{\sim} \chi^2_{a_n'}$, where $r_K' = r_K P(m \in [a, b])$ and $a_n' = s_K(b - a)/h$. The last result and Equation (C.1) imply that

$$r_K T_n = r_K' T_n' + O_P(n^{-1/2}h^{-1}) \overset{a}{\sim} \chi^2_{a_n}.$$

[Received March 2008. Revised February 2009.]

REFERENCES

Aït-Sahalia, Y., and Duarte, J. (2003), "Nonparametric Option Pricing Under Shape Restrictions," *Journal of Econometrics*, 116, 9–47.

Aït-Sahalia, Y., and Lo, A. W. (1998), "Nonparametric Estimation of State-Price Densities Implicit in Financial Assets Prices," *Journal of Finance*, 53, 499–548.

Amin, K., and Ng, V. (1993), "Option Valuation With Systematic Stochastic Volatility," *Journal of Finance*, 48, 881–910.

Bakshi, G., and Chen, Z. (1997), "An Alternative Valuation Model for Contingent Claims," *Journal of Financial Economics*, 44, 123–165.

Bakshi, G., Cao, C., and Chen, Z. (1997), "Empirical Performance of Alternative Option Pricing Models," *Journal of Finance*, 52, 2003–2049.

Barone-Adesi, G., Engle, R. F., and Mancini, L. (2008), "A GARCH Option Pricing Model With Filtered Historical Simulation," *Review of Financial Studies*, 21, 1223–1258.

Bates, D. S. (1991), "The Crash of '87: Was It Expected? The Evidence From Options Markets," *Journal of Finance*, 46, 1009–1044.

——— (1996), "Jumps and Stochastic Volatility: Exchange Rate Processes Implicit in Deutsche Mark Options," *Review of Financial Studies*, 9, 69–107.

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

Bollerslev, T. P. (1986), "Generalized Autoregressive Conditional Heteroskedasticity," *Journal of Econometrics*, 31, 307–327.

Bollerslev, T. P., Chou, R. Y., and Kroner, K. F. (1992), "ARCH Modelling in Finance: A Review of the Theory and Empirical Evidence," *Journal of Econometrics*, 52, 5–59.

Bossaerts, P., and Hillion, P. (1997), "Local Parametric Analysis of Hedging in Discrete Time," *Journal of Econometrics*, 81, 243–272.

——— (2003), "Local Parametric Analysis of Derivatives Pricing and Hedging," *Journal of Futures Markets*, 6, 573–605.

Breeden, D., and Litzenberger, R. (1978), "Prices of State Contingent Claims Implicit in Option Prices," *Journal of Business*, 51, 621–657.

Carr, P., Geman, H., Madan, D. B., and Yor, M. (2003), "Stochastic Volatility for Lévy Processes," *Mathematical Finance*, 13, 345–382.

Chernov, M., and Ghysels, E. (2000), "A Study Towards a Unified Approach to the Joint Estimation of Objective and Risk Neutral Measures for the Purpose of Options Valuation," *Journal of Financial Economics*, 56, 407–458.

Christoffersen, P., Heston, S., and Jacobs, K. (2006), "Option Valuation With Conditional Skewness," *Journal of Econometrics*, 131, 253–284.

Cont, R., and Tankov, P. (2004), "Non-Parametric Calibration of Jump-Diffusion Option Pricing Models," *Journal of Computational Finance*, 7, 1–49.

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

Duan, J.-C. (1995), "The GARCH Option Pricing Model," *Mathematical Finance*, 5, 13–32.

Dumas, B., Fleming, J., and Whaley, R. E. (1998), "Implied Volatility Functions: Empirical Tests," *Journal of Finance*, 53, 2059–2106.

Dupire, B. (1994), "Pricing With a Smile," *Risk*, 7, 18–20.

Engle, R. F. (1982), "Autoregressive Conditional Heteroskedasticity With Estimate of the Variance of United Kingdom Inflation," *Econometrica*, 50, 987–1007.

Fan, J. (1992), "Design-Adaptive Nonparametric Regression," *Journal of the American Statistical Association*, 87, 998–1004.

Fan, J., and Gijbels, I. (1995), "Data-Driven Bandwidth Selection in Local Polynomial Fitting: Variable Bandwidth and Spatial Adaptation," *Journal of the Royal Statistical Society, Ser. B*, 57, 371–394.

Fan, J., and Jiang, J. (2007), "Nonparametric Inference With Generalized Likelihood Ratio Tests" (with discussion), *Test*, 16, 409–478.

Fan, J., and Yao, Q. (2003), *Nonlinear Time Series: Nonparametric and Parametric Methods*, New York: Springer-Verlag.

Fan, J., Zhang, C. M., and Zhang, J. (2001), "Generalized Likelihood Ratio (GLR) Statistics and Wilks Phenomenon," *The Annals of Statistics*, 29, 153–193.

Fan, Y., and Ullah, A. (1999), "Asymptotic Normality of a Combined Regression Estimator," *Journal of Multivariate Analysis*, 71, 191–240.

Fleming, J., Ostdiek, B., and Whaley, R. E. (1996), "Trading Costs and the Relative Rates of Price Discovery in the Stock, Futures, and Option Markets," *Journal of Futures Markets*, 16, 353–387.

Gagliardini, P., Gourieroux, C., and Renault, E. (2005), "Efficient Derivative Pricing by Extended Method of Moments," Working Paper 2005-05, University of St. Gallen.

Galai, D. (1983), "The Components of the Return From Hedging Options," *Journal of Business*, 56, 45–54.

Garcia, R., and Gençay, R. (2000), "Pricing and Hedging Derivative Securities With Neural Networks and a Homogeneity Hint," *Journal of Econometrics*, 94, 93–115.

Ghysels, E., Harvey, A. C., and Renault, E. (1996), "Stochastic Volatility," in *Handbook of Statistics*, eds. G. S. Maddala and C. R. Rao, Amsterdam: North-Holland, pp. 119–191.

Giacomini, R., Gottschling, A., Haefke, C., and White, H. (2008), "Mixtures of *t*-Distributions for Finance and Forecasting," *Journal of Econometrics*, 144, 175–192.

Glad, I. K. (1998), "Parametrically Guided Non-Parametric Regression," *Scandinavian Journal of Statistics*, 25, 649–668.

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

Heston, S. (1993), "A Closed-Form Solution for Options With Stochastic Volatility, With Applications to Bond and Currency Options," *Review of Financial Studies*, 6, 327–343.

Heston, S., and Nandi, S. (2000), "A Closed-Form GARCH Option Valuation Model," *Review of Financial Studies*, 13, 585–625.

Hjort, N. L., and Glad, I. K. (1995), "Nonparametric Density Estimation With a Parametric Start," *The Annals of Statistics*, 23, 882–904.

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

Jones, C. S. (2003), "The Dynamics of Stochastic Volatility: Evidence From Underlying and Options Markets," *Journal of Econometrics*, 116, 181–224.

Madan, D. B., Carr, P., and Chang, E. (1998), "The Variance Gamma Process and Option Pricing," *European Finance Review*, 2, 79–105.

Press, W. H., and Tukey, J. W. (1956), "Power Spectral Methods of Analysis and Their Application to Problems in Airplane Dynamics," in *Bell Telephone System Monograph*, Vol. 2606, New York: Bell Telephone System.

Renault, E., and Touzi, N. (1996), "Option Hedging and Implicit Volatilities in a Stochastic Volatility Model," *Mathematical Finance*, 6, 279–302.

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

Ruppert, D., Sheather, S. J., and Wand, M. P. (1995), "An Effective Bandwidth Selector for Local Least Squares Regression," *Journal of the American Statistical Association*, 90, 1257–1270.

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

Scott, L. O. (1997), "Pricing Stock Options in a Jump-Diffusion Model With Stochastic Volatility and Interest Rates: Application of Fourier Inversion Methods," *Mathematical Finance*, 7, 413–426.

Zhang, C. M. (2003), "Calibrating the Degrees of Freedom for Automatic Data-Smoothing and Effective Curve-Checking," *Journal of the American Statistical Association*, 98, 609–628.