

Option Pricing with Aggregation of Physical Models and Nonparametric Statistical Learning*

Jianqing Fan^a

Loriano Mancini^b

^aBendheim Center for Finance and
Department of Operations Research and Financial Engineering
Princeton University, USA

^bSwiss Banking Institute, University of Zurich, Switzerland

9th December 2006

Keywords: Nonparametric regression, state price distribution, out-of-sample analysis, hedge ratio, model misspecification.

JEL Classifications: C14, G13.

***Correspondence Information:** Jianqing Fan, ORFE, E-Quad, Princeton University, Princeton, NJ 08544. E-mail: jqfan@princeton.edu. E-mail address for Loriano Mancini: mancini@isb.unizh.ch. For helpful comments, we thank Giovanni Barone-Adesi, Eric Ghysels, Rajna Gibson and seminar participants at the conferences “Financial Engineering and Risk Management” 2006, Xiamen University, “Far Eastern Meeting of Econometrics Society” 2006, Beijing, “Statistics at the Frontier of Science” 2006, Banff International Research Station, “International Statistics Forum” 2006, Beijing, the University of Minnesota, the Georgia Institute of Technology, the Waseda University, the Wharton School and the London School of Economics. Fan’s research was supported by the National Science Foundation grant DMS-0532370. 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.

Option Pricing with Aggregation of Physical Models and Nonparametric Statistical Learning

Abstract

Financial models are largely used in option pricing. These physical models capture several salient features of asset price dynamics. The pricing performance can be significantly enhanced when they are combined with nonparametric learning approaches, that empirically learn and correct pricing errors through estimating state price distributions. In this paper, we propose a new semiparametric method for estimating state price distributions and pricing financial derivatives. This method is based on a physical model guided nonparametric approach to estimate the state price distribution of a normalized state variable, called the Automatic Correction of Errors (ACE) in pricing formulae. Our method is easy to implement and can be combined with any model based pricing formula to correct the systematic biases of pricing errors and enhance the predictive power. 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.

Introduction

Over the last three decades, there have been substantial efforts in extending the Black and Scholes (1973) model along several directions. These efforts aim at developing more flexible physical dynamics of asset prices leading to more accurate option pricing formulae. Examples include the jump-diffusion models of Bates (1991) and Madan, Carr, and Chang (1998), the stochastic volatility models of Hull and White (1987), Heston (1993), and Melino and Turnbull (1995), the stochastic volatility and stochastic interest rates models of Amin and Ng (1993), 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 physical price movements more plausible.

For instance, Bakshi, Cao, and Chen (1997) derived a general closed-form pricing formula for a family of jump-diffusion stochastic volatility models with stochastic interest rates. Essentially, these models produce a pricing formula which depends on option characteristics such as spot stock price, volatility, time to maturity, risk-free interest rate, and dividend rate. The larger the family of models, the more flexible the pricing formula is. These models encompass many commonly used ones in practice. The parameters in the model are then calibrated to fit the observed option prices. In a similar effort but a different direction, Duan (1995), Heston and Nandi (2000), and Barone-Adesi, Engle, and Mancini (2006) take advantage of the flexibility of the GARCH models to provide an option pricing formula. By assuming that the asset price dynamics under the risk neutral measure follow a GARCH model, the pricing formula can be analytically derived (Heston and Nandi (2000)) or numerically computed via statistical simulations (Barone-Adesi, Engle, and Mancini (2006)). The parameters in the GARCH model are then calibrated to best fit the observed option prices.

The aforementioned parametric models attempt to capture certain salient features of observed price dynamics. However, these models cannot be derived from comprehensive economic theories, often rely on assumptions concerning the risk neutral asset dynamics, and need to be simple and convenient to allow for the derivation of pricing formulae. Hence these models cannot be expected to capture all the relevant features of the involved pricing mechanisms. Indeed, there are always limitations on the performance of these physical modeling techniques and model misspecification is a major concern that can lead to erroneous valuation and hedging strategies. Despite the aforementioned concerns, these pricing formulae have been proven useful for several practices. For instance, even though these models might not be correct, traders use those formulae to obtain initial reference prices. When they are collectively used by many practitioners, these formulae become naturally a good first order approximation of option prices—although several attempts have been made to enhance their practical utilities.

A well documented empirical feature of implied volatilities is the asymmetry in the volatility smiles. These smiles can be induced by the negative correlation between asset returns and volatilities (see, for instance, Renault and Touzi (1996)), and the negative skewness of return innovations (see Barone-Adesi, Engle, and Mancini (2006)).¹ Sensible option pricing models should account for such volatility smiles and an effective

¹In a recent review Sundaresan (2000) remarks that the term structure of implied volatilities is still puzzling in stochastic volatility models. For instance Comte and Renault (1998) attempt to explain the term structure of volatility smiles using long

empirical approach is the ad hoc Black-Scholes model introduced by Dumas, Fleming, and Whaley (1998). The method is to fit a quadratic function to the implied volatilities and then to price options by plugging the estimated volatilities in the Black-Scholes formula. Dumas, Fleming, and Whaley (1998) show that this approach outperforms the deterministic volatility function models by Derman and Kani (1994), Dupire (1994), and Rubinstein (1994). These results show that empirical approaches can be combined with model based pricing formulae to enhance practical utilities of pricing formulae.

In this paper, instead of attempting to improve option pricing formulae by introducing even more flexible option pricing models, we propose a method of improvement in an orthogonal direction. Our approach to price options is based on the nonparametric correction of pricing errors of a pricing model. Given a pricing formula derived from a physical model, for each relevant maturity we calibrate the model parameters to best fit the observed option prices. Then we nonparametrically learn the pricing errors induced by the pricing formula and correct them. This is achieved via a nonparametric estimate of the state price distribution using the guidance of the parametric model on the state price distribution as initial estimate. Fitting a separate curve for each given time to maturity gives us the flexibility to model individualized pricing error functions, and hence reduces the biases in the empirical fitting. The approach will take into account that only a limited number of options is traded for each maturity.

We use a nonparametric method to correct the pricing errors as the forms of pricing errors are hard to determine, varying over time and time to maturity. Nonparametric methods have the flexibility to discover the nonlinear relation between pricing errors and moneyness. For each given time to maturity and other option characteristics, we estimate the state price distribution, that is the integral of the state price density; see for instance Cox and Ross (1976) and Harrison and Kreps (1979). In the nonparametric literature (see for e.g. Fan and Yao (2003)), it is well known that the distribution function is much easier to estimate, admitting a faster rate of convergence, than the density function. Hence we estimate the state price distribution instead of the state price density. This is another important aspect of our methodological contribution to derivative pricing.

The state price distribution is related to the expected pay-off of a tradable portfolio, consisting of two positions in call options (see Figure 1). Hence the problem of estimating the state price distribution becomes a nonparametric regression problem. The state price distribution² is always decreasing, has a water fall shape, and is in a neighborhood of a model based pricing formula (see Figure 2). Simple and direct application of nonparametric regression does not take advantage of this knowledge. Indeed, the distribution function admits different degrees of smoothness³ and nonparametric approaches do not work well without the use of involved variable smoothing techniques. This issue will be demonstrated in the empirical study. Our idea is first to estimate the main shape of the state price distribution by using a pricing formula derived from a sensible physical model and then estimate and correct the pricing errors using a nonparametric approach. This method, called the Automatic Correction of Errors (ACE) of a pricing formula, will reduce substantially

memory, fractional integrated volatility processes. These methods are more involved than the approach proposed here.

²More precisely, it is one minus the state price distribution, which is called the survivor function in statistics and risk analysis.

³The first derivative of the function is much larger in the middle than at both ends.

the pricing errors.

Our approach can be regarded as the aggregation of physical model based prices with statistical learning of pricing errors, estimated and corrected via a nonparametric approach. The nonparametric learning and correction of pricing errors are very easy to implement. This method can be combined with any model based pricing formula. The particular one that we use in the paper is inspired by the ad hoc Black-Scholes model, as it is frequently employed in financial industry. This is a major advantage of our method. The approach provides an arbitrage-free method for pricing other securities. The state price distribution is estimated from “fundamental” liquid option prices and can be used to price other less liquid, such as over-the-counter derivatives and more complex or nontraded options.⁴

In the implementation, we will use the ad hoc Black-Scholes method as an initial estimate of the state price distribution. The ad hoc Black-Scholes method is very easy to implement and fast to compute. As a result, our aggregated method is also very fast to compute and implement. In fact, it is many orders of magnitude faster than calibration based approaches such as Duan (1995), Bakshi, Cao, and Chen (1997), Heston and Nandi (2000), and Barone-Adesi, Engle, and Mancini (2006). We also investigate a new approach that combines the ad hoc Black-Scholes method with a nonparametric estimate of the implied volatility curve for each relevant maturity. This semiparametric Black-Scholes model has the flexibility to fit an implied volatility curve for each given time to maturity, and hence reduces the biases in the empirical pricing. It also takes into account the discreteness of traded time to maturities. This approach is inspired by the nonparametric method of Aït-Sahalia and Lo (1998), who fit two-dimensional functions to the implied volatility by a different nonparametric function. Compared to their method, however, this approach is easier to implement. In addition, instead of pulling the information from option prices with different maturities on a given day—that are very discrete—we aggregate the information from options with the same maturities but traded on consecutive dates.

In the empirical application we consider European options on the S&P 500 index traded from January 2002 to December 2004. We compare our ACE method to several alternative methods. The first method is the ad hoc Black-Scholes model of Dumas, Fleming, and Whaley (1998), that provides an interesting and challenging benchmark as it allows for different implied volatilities to price different options. The second method is the semiparametric Black-Scholes model inspired by Aït-Sahalia and Lo (1998), that is expected to outperform the ad hoc Black-Scholes model. The third method is the GARCH option pricing model introduced by Heston and Nandi (2000). This parametric method allows us to evaluate the relative performance of nonparametric and parametric approaches. The fourth method is the nonparametric regression approach directly applied to estimate the state price distribution. We show that our aggregated ACE method outperforms both physical model based approaches (the ad hoc Black-Scholes and the calibrated GARCH models) and nonparametric and semiparametric methods, in terms of fitting and prediction of option prices as well as hedging performance. Furthermore, we develop a formal nonparametric test, using the generalized likelihood ratio test of Fan, Zhang, and Zhang (2001), to show that the nonparametric correction in the ACE approach is effective in reducing the

⁴It is well-known that markets have to be dynamically complete for such prices to be meaningful; see for instance Constantinides (1982) and Duffie and Huang (1985). This assumption is usually adopted in derivative pricing models and we also adopt it here.

pricing errors. These two pieces of empirical results provide stark evidence on the power of the nonparametric learning of pricing errors in option pricing and hedging.

The paper is organized as follows. Section 1 introduces the state price distribution, the ad hoc Black-Scholes and the direct nonparametric methods. It also presents our ACE pricing method and the nonparametric evaluation test. Section 2 recalls the semiparametric Black-Scholes and the GARCH pricing models. Section 3 presents the empirical analysis, that is in-sample and out-of-sample pricing, and hedging performances of the previous option pricing models and Section 4 concludes.

1 Estimation of the state price distribution

The fundamental idea of our approach is to estimate the cumulative state price distribution via traded portfolios (see Figure 1). Let S_t be the price at time t of the underlying asset of option contracts. The state price density $f^*(\cdot)$ is the conditional density of the asset price S_T at maturity T given S_t under the risk neutral measure (Cox and Ross (1976), Harrison and Kreps (1979)), and it has a simple relation with the price of the option written on the asset. Let C_t denote the call price at time t with strike price X , and time to maturity $\tau = T - t$. Then

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

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

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

where $\bar{F}^*(x) = 1 - F^*(x)$, is the state price survivor function of S_T . **JQ: I added the following “emphasized” sentences and the footnote:** *In general, the (forward) price of a contingent claim with pay-off function $\psi(S_T)$ can be easily expressed in terms of the state price survivor function,*

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

The pay-off function, ψ , needs to satisfy some mild regularity conditions⁵ that are usually verified by derivative contracts traded on the market. Hence for the pricing purposes knowing the state price survivor function is equivalent to knowing the state price density. In the following for simplicity we consider only the pricing of call options. The pricing of other derivative contracts can be similarly obtained. By no arbitrage S_T is equal to the futures price of the asset at time T . Let

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

be the futures price of the asset at time t , where $\delta_{t,\tau}$ is the dividend yield paid by the asset between t and $T = t + \tau$. We denote the futures price always using both time subscripts, $F_{t,\tau}$, to avoid any confusion with

⁵The pay-off function, ψ , needs to be bounded at zero, i.e. $\psi(0) < \infty$, not increasing too rapidly, i.e. $\lim_{y \rightarrow \infty} \psi(y) \bar{F}^*(y) = 0$ and the integral in the right hand side has to be well-defined, and being sufficiently smooth, i.e. $\psi'(y) < \infty$.

state price distribution functions. By the change of variable

$$C_t = e^{-r_t, \tau} F_{t, \tau} \int_{m_t}^{\infty} \bar{F}(u) du, \quad (1)$$

where $m_t = X/F_{t, \tau}$ is the moneyness and $\bar{F}(u) = 1 - F^*(F_{t, \tau}u)$ is the state price survivor function in the normalized scale,⁶ that is in the last integral the futures price at time t , $F_{t, \tau}$, is normalized to \$1. In the sequel we also denote the state price survivor function by \bar{F}_t to emphasize the dependence on the information at time t . As an example in the Black-Scholes model the state price distribution is the log-normal distribution and by the change of variable,

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

where σ is the constant volatility and Φ the standard Gaussian distribution function. In general, state price densities and distributions have no closed-form expressions and numerical procedures are usually adopted. For instance, in jump diffusion models with nonparametric Lévy measure Cont and Tankov (2004) suggest to approximate the state price density by using a mixture of log-normal distributions.

Using equation (1), pricing European options reduces to the estimation of the state price survivor function \bar{F} . The parametric approach amounts to assuming a risk neutral dynamic for the underlying asset, and then deriving the form of the state price distribution. For example, the survivor function \bar{F} in (2) is derived under the Black-Scholes model and depends only on a few parameters. The pricing formula for the GARCH model derived by Heston and Nandi (2000), though not as explicit as the Black-Scholes formula, implies a certain parametric form for the state price distribution. Then the parametric approach infers the model parameters from traded options.

Our approach is nonparametric. It directly infers the state price survivor function from traded options with different moneyness for each given time to maturity τ . 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 avoids the danger of model misspecification and allows for fast estimation of the state price distribution. In addition, our nonparametric method takes advantage of the explanatory power of parametric pricing formulae and searches for pricing formulae around the parametric ones. Combining the power of model based approaches and nonparametric learning is another important aspect of our methodological contribution.

1.1 Traded options and state price distribution

We discuss how to infer the state price distribution from traded options with different moneyness; see Figure 1. Let $X_1 < X_2$ be two consecutive strike prices of traded options. 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 pay-off function close to a digital call option pay-off with indicator function, $I(x > (X_1 + X_2)/2)$, which takes values one if x

⁶This normalized approach has the advantage not only in understanding the scale of the strike price X , but also in accounting for the dividends paid by the stock.

exceeds $(X_1 + X_2)/2$ and zero otherwise. Hence

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

where \approx means approximately equal, the expectation is under the risk neutral measure and $C(X_i)$ is the European call option price with strike X_i at time t , omitting the time subscripts t and τ . **JQ: I added the following “emphasized” sentences:** *The accuracy of this approximation will be reflected in the pricing performance of our approach based on the state price distribution. Eventually, whether or not this approximation is sufficiently accurate for pricing purposes is an empirical issue that will be extensively investigated in Section 3.* As derived in Appendix A, the mid-point gives the best approximation in terms of the order of approximation error. We summarize the theoretical findings in the following proposition.

Proposition 1 *Let $C(X)$ be the price of the European call option with strike price X . For two given strike prices $X_1 < X_2$, we have⁷*

$$e^{r\tau} \frac{C(X_1) - C(X_2)}{X_2 - X_1} = \bar{F}(\bar{m}) + O((m_1 - m_2)^2), \quad (3)$$

where $m_i = X_i/F_{t,\tau}$ ($i = 1, 2$) is the corresponding moneyness and $\bar{m} = (m_1 + m_2)/2$. The approximation error is bounded by

$$-\frac{1}{24} \left\{ \min_{m_1 \leq \xi \leq m_2} f'(\xi) \right\} (m_2 - m_1)^3,$$

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

The proof is given in Appendix A. Proposition 1 provides a theoretical basis for inferring the state price distribution from traded options. One advantage of using the futures price $F_{t,\tau}$ is that it accounts for the dividends paid by the asset, the risk-free rate, and the current spot asset price. To simplify the notation, let $C_{t,i} = C_t(X_{t,i})$ denote the call option price with strike price $X_{t,i}$ or moneyness $m_{t,i}$ at time t . Let us order the moneyness $\{m_{t,i}\}$ traded at time t in an ascending order. Denote by

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

Then according to equation (3) we have

$$Y_{t,i} = \bar{F}(\bar{m}_{t,i}) + \varepsilon_{t,i}, \quad (4)$$

where $\varepsilon_{t,i}$ is the idiosyncratic noise. This approach reduces the option pricing problem to the nonparametric regression problem. Hence a nonparametric technique can be used to estimate \bar{F} . In contrast to other nonparametric methods such as those in Ait-Sahalia and Lo (1998) and Ait-Sahalia and Duarte (2003), our method nonparametrically estimates the state price survivor function rather than the state price density. Equation (3) shows that the former, $\bar{F}(m)$, is almost directly observable on the option market, while the state price density is not and has to be recovered for instance taking the second derivative of the call option function with respect to the strike price; see for example Ait-Sahalia and Lo (1998) and the references therein.

⁷We have assumed that these two options have the same option characteristics, such as time to maturity, except the strike price. For clarity of exposition we have suppressed the dependence of the option prices on other characteristics.

Moreover, the state price distribution is much easier to estimate, admitting a faster rate of convergence than the state price density (Fan and Yao (2003)). In our empirical studies, we use closing option prices to estimate the state price distribution function for each relevant maturity. This procedure gives about 30–40 data points for each day t . Therefore, as for e.g. in Ait-Sahalia and Lo (1998), we aggregate the data around a given time point t_0 to exploit the continuity of the option pricing formula as a function of time.

The particular nonparametric method that we will use is the local linear regression. It has several advantages, including automatic boundary correction, high statistical efficiency, and easy bandwidth selection, as demonstrated in Fan (1992) and Fan and Gijbels (1995). For an overview of the local linear estimator and other related techniques, we refer the reader to Fan and Yao (2003). At a given time t_0 , the direct nonparametric estimator of \bar{F} with maturity τ and moneyness m is given by the time-weighted local linear regression

$$\min_{\beta_0, \beta_1 \in \mathbb{R}^2} \sum_{t=t_0-d}^{t_0+d} \lambda^{|t_0-t|} \sum_{i=1}^{N_t} (Y_{t,i} - \beta_0 - \beta_1 (\bar{m}_{t,i} - m))^2 K_h(\bar{m}_{t,i} - m), \quad (5)$$

where $\lambda \in (0, 1]$ is a smoothing parameter in time, K is a kernel function, h is the bandwidth used to fit a local linear model, $K_h(u) = h^{-1}K(u/h)$, and $N_t + 1$ is the number of options traded at time t for the maturity of interest. Denoting by $\hat{\beta}_0$ and $\hat{\beta}_1$ the resulting minimizers, $\hat{F}(m) = \hat{\beta}_0$ is the direct nonparametric estimate of the state price survivor function. With the estimated \hat{F} , the call option prices are computed via equation (1). The first summation in (5) aggregates information from options traded on consecutive dates to exploit the continuity of the state price distribution as a function of time. In our implementation we set $d = 2$, that is we use a week of daily closing prices to infer the state price distribution and achieving a reasonable sample size of 150–200 data points. Options traded on different days have slightly different time to maturities and the time-weight $\lambda^{|t_0-t|}$ accounts for this effect. Notice that for fitting purposes on day t_0 , we use the data up to time $t_0 + d$, borrowing future information. An alternative approach is to use the data only up to date t_0 , and in this case the first summation is from $t_0 - 2d$ to t_0 . We also implemented this second approach and will discuss the results in Section 3. The second summation in (5) 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 kernel⁸ $K(x) = (1 - x^2)^+$ that vanishes outside the interval $(-1, 1)$, the local linear regression in (5) uses only the data with moneyness $\bar{m}_{t,i}$ in the interval $m \pm h$. Hence the bandwidth controls the effective sample size. As an example, Figure 2 shows the nonparametric estimate of the survivor function \bar{F} when t_0 is equal to December 29, 2004 and $Y_{t,i}$'s are observed on December 27–31, 2004 with maturities 169–173 days. Visual inspection of the fitting shows that the direct nonparametric method does a reasonably good job. However, even small errors in estimating the state price distribution can translate into large pricing errors. This is particularly true for the moneyness around one as demonstrated

⁸This kernel is optimal in terms of minimizing the mean square error of the resulting nonparametric estimator. Here the superscript “+” denotes the positive part of a variable: $x^+ = x$ if $x > 0$ and zero otherwise.

in Figure 3. The unsatisfactory pricing errors of the direct nonparametric approach are mainly due to the lack of use of the prior knowledge on the overall shape of the survivor curve. To exploit the main shape of the survivor function, we first need to give a crude estimate of the shape and then correct the difference by a nonparametric method. The particular method that we use is inspired by the ad hoc Black-Scholes method as it is very simple to implement and to compute. Other parametric and possibly more involved models, such as those investigated by Bakshi, Cao, and Chen (1997) could be used. However, given the subsequent nonparametric estimation of pricing errors it is advisable to use a simple parametric model in the first step.

1.2 Ad hoc Black-Scholes model

A well documented empirical feature of implied volatilities is the so called volatility smile; see Figure 4. To account for this phenomenon and provide a benchmark option pricing model, Dumas, Fleming, and Whaley (1998) introduce an ad hoc Black-Scholes model where the implied volatilities, σ^{bs} , are smoothed across moneyness, m , by fitting a parabolic function

$$\sigma_{t,i}^{\text{bs}} = a_0 + a_1 m_{t,i} + a_2 m_{t,i}^2 + \text{error}_{t,i}, \quad (6)$$

where $\sigma_{t,i}^{\text{bs}}$ denotes the implied volatility observed on day t for a given maturity T and moneyness $m_{t,i}$, $t = t_0 - d, \dots, t_0 + d$, and $i = 1, \dots, N_t + 1$. Implied volatilities observed on different days have slightly different time to maturities, $\tau = T - t$, for $t = t_0 - d, \dots, t_0 + d$. In our empirical application, on each day t_0 and for each maturity T the quadratic function (6) is estimated using time-weighted least square regressions with time-weight $\lambda^{|t_0-t|}$, that is down weighting implied volatilities not observed on day t_0 . The fitted value $\hat{\sigma}^{\text{bs}}$ is used to price options via the Black-Scholes model, that is plugging $\hat{\sigma}^{\text{bs}}$ in the Black-Scholes formula, $C_{t,i}^{\text{bs}} = e^{-r\tau} (F_{t,\tau} \Phi(d_1) - X_{t,i} \Phi(d_2))$, where $d_1 = (\log(F_{t,\tau}/X_{t,i}) + \sigma^2\tau/2) / (\sigma\sqrt{\tau})$, $d_2 = d_1 - \sigma\sqrt{\tau}$, and setting $\sigma = \hat{\sigma}^{\text{bs}}$. We shall use the estimated coefficients, \hat{a}_j , $j = 0, 1, 2$, also to implement our approach; see equation (7). A linear term in the time to maturity, $a_3 \tau$, could be added to the regression model (6), but to avoid overfitting we use the simple model (6) as our benchmark. Indeed, Dumas, Fleming, and Whaley (1998, Tables III and IV) show that the model (6) tends to outperform the “extended” model with τ regressor in the out-of-sample and hedging exercises. As an example, Figure 4 shows the implied volatilities of the call option prices analyzed in the previous section and observed on the week December 27–31, 2004, with time to maturities from 173 to 169 days. The fitted $\hat{\sigma}^{\text{bs}}$ represents the estimated volatilities for the moneyness observed on December 29, 2004. Some degrees of lack of fit are evidenced due to the inflexibility of the quadratic form. This translates into systematic pricing errors, as demonstrated in Figure 3.

Although theoretically inconsistent, ad hoc Black-Scholes methods are routinely used in the option pricing industry and they represent a challenging benchmark as they allow for different implied volatilities to price different options. Moreover, Dumas, Fleming, and Whaley (1998) show that this approach outperforms the deterministic volatility function option valuation model introduced by Derman and Kani (1994), Dupire (1994), and Rubinstein (1994).

1.3 Nonparametric learning of pricing errors

A weakness of the direct nonparametric approach (5) is that a constant bandwidth h is used over the whole domain of \bar{F} in regression (4). As shown in Figure 3, the pricing errors of this approach tend to be larger when the survivor function is steeper ($m \approx 1$) than when the survivor function is flatter ($m \ll 1$ and $m \gg 1$). In addition, the direct nonparametric approach does not explicitly take into account the implied volatility smile that is a persistent phenomenon in option markets; see for instance Renault and Touzi (1996). We propose to address these aspects by the following 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 (2) and the ad hoc Black-Scholes model. For a given maturity, T , to account for the slightly different time to maturities during the week or the $2d + 1$ trading days, we introduce a time dependent calibration parameter ϑ_t . The proposed survivor function is

$$\bar{F}_{\text{LN}}(m; \vartheta_t) = 1 - \Phi \left[\frac{\log(m) + ((a_0 + a_1 m + a_2 m^2) \vartheta_t)^2 / 2}{(a_0 + a_1 m + a_2 m^2) \vartheta_t} \right], \quad (7)$$

where the coefficients a_j , $j = 0, 1, 2$, are estimated using the implied volatilities observed on the $2d + 1$ days as in the ad hoc Black-Scholes model (6). The choice of the log-normal distribution is motivated by the Black-Scholes model under which the implied volatilities are computed. The calibration parameter ϑ_t is determined by minimizing the distance between empirical and theoretical values:

$$\hat{\vartheta}_t = \arg \min_{\vartheta \in \mathbb{R}} \sum_{i=1}^{N_t} (Y_{t,i} - \bar{F}_{\text{LN}}(\bar{m}_{t,i}; \vartheta))^2. \quad (8)$$

The resulting parametric estimate of the state price survivor function, $\bar{F}_{\text{LN}}(m; \hat{\vartheta}_t)$, combines the log-normal survivor function and the ad hoc Black-Scholes method. Notice that $\bar{F}_{\text{LN}}(m; \vartheta_t)$ reduces to the Black-Scholes log-normal survivor function 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, $\bar{F}_{\text{LN}}(m; \hat{\vartheta}_t)$. Solving the minimization problem (8) amounts to price the forward digital call options, $Y_{t,i}$'s, using the physical model, $\bar{F}_{\text{LN}}(m; \hat{\vartheta}_t)$. The pricing errors, $\tilde{Y}_{t,i} = Y_{t,i} - \bar{F}_{\text{LN}}(\bar{m}_{t,i}; \hat{\vartheta}_t)$, induced by the model $\bar{F}_{\text{LN}}(m; \hat{\vartheta}_t)$ will be empirically learned and corrected by a nonparametric term, $\bar{F}_{t,c}(m)$, to improve the pricing performance. The nonparametric learning on the correction term, $\bar{F}_{t,c}(m)$, can be achieved by the local linear fit to the data, $\{(\bar{m}_{t,i}, \tilde{Y}_{t,i}), i = 1, \dots, N_t, t - t_0 \in [-d, d]\}$. Similarly to the direct nonparametric approach (5) the correction term, $\bar{F}_{t,c}(m)$, is estimated using the time-weighted nonparametric regression

$$\min_{\beta_0, \beta_1 \in \mathbb{R}^2} \sum_{t=t_0-d}^{t_0+d} \lambda^{|t_0-t|} \sum_{i=1}^{N_t} \left(\tilde{Y}_{t,i} - \beta_0 - \beta_1 (\bar{m}_{t,i} - m) \right)^2 K_h(\bar{m}_{t,i} - m), \quad (9)$$

where $Y_{t,i}$ is now replaced by $\tilde{Y}_{t,i}$. Denoting by $\tilde{\beta}_0$ and $\tilde{\beta}_1$ the resulting minimizers, the nonparametric correction term at the point m is given by $\hat{\bar{F}}_{t,c}(m) = \tilde{\beta}_0$. The nonparametric regression is now applied to the $\tilde{Y}_{t,i}$'s which are expected to be more spatially homogeneous than the $Y_{t,i}$'s with approximately the same degree of smoothness. Hence the constant bandwidth, h , should be a reasonable choice for estimating the

correction function, $\bar{F}_{t,c}(m)$. The shape of $\bar{F}_{t,c}(m)$ is usually not easy to determine. Since the function can vary over time a nonparametric learning technique is particularly appealing and it is also easy to implement. **JQ: I added the following “emphasized” sentences:** *An alternative approach would be to investigate the temporal dynamic of $\bar{F}_{t,c}$, trying to learn and correct the pricing errors in a Bayesian fashion. However, this approach would be more difficult to implement and computationally more intensive than our nonparametric approach.* Hence the state price survivor function, $\bar{F}_t(m)$, can be represented as

$$\bar{F}_t(m) = \bar{F}_{\text{LN}}(m; \hat{\vartheta}_t) + \bar{F}_{t,c}(m), \quad (10)$$

where $\bar{F}_{\text{LN}}(m; \hat{\vartheta}_t)$ is the parametric fit of the state price survivor function and $\bar{F}_{t,c}(m)$ the nonparametric correction of the pricing errors induced by $\bar{F}_{\text{LN}}(m; \hat{\vartheta}_t)$. Interestingly, any state price survivor function can be represented as in equation (10). Intuitively, replacing $\bar{F}_{t,c}(m) \approx Y_t - F_{\text{LN}}(m; \hat{\theta}_t)$ in equation (10), the parametric terms, $F_{\text{LN}}(m; \hat{\theta}_t)$, cancel out and this shows that the estimation of the state price survivor function is essentially nonparametric. Hence the nonparametric estimate of $\bar{F}_{t,c}(m)$ is equivalent to the nonparametric estimate of $\bar{F}_t(m)$ and equation (10) provides a flexible representation for any the state price survivor function. We point out that $\bar{F}_{t,c}$ is not a survivor function, but we use a similar notation to emphasize that it is the correction term of the parametric survivor function, \bar{F}_{LN} . When the true state price survivor function is exactly given by \bar{F}_{LN} or any other parametric model, the correction term $\bar{F}_{t,c}(m)$ has to be zero—but of course this is rarely the case in practice. Equations (4) and (10) imply that

$$Y_{t,i} = \bar{F}_{\text{LN}}(\bar{m}_{t,i}; \hat{\vartheta}_t) + \bar{F}_{t,c}(\bar{m}_{t,i}) + \varepsilon_{t,i} \quad (11)$$

and this equation shows that $\bar{F}_{t,c}(m)$ is the regression function for the data, $\tilde{Y}_{t,i} = Y_{t,i} - \bar{F}_{\text{LN}}(\bar{m}_{t,i}; \hat{\vartheta}_t)$, on the moneyness, $\bar{m}_{t,i}$. The call price of the European option with moneyness m_t and time to maturity τ , according to (1), can be summarized in the following proposition.

Proposition 2 *The price of the European call option with moneyness m_t and time to maturity τ can be decomposed as*

$$C_t = e^{-r\tau} F_{t,\tau} \int_{m_t}^{\infty} \bar{F}_{\text{LN}}(u; \hat{\vartheta}_t) du + e^{-r\tau} F_{t,\tau} \int_{m_t}^{\infty} \bar{F}_{t,c}(u) du. \quad (12)$$

The first part is the option pricing formula derived from the physical model for the asset dynamic under the risk neutral measure and the second part is the correction of the pricing error due to misspecification of the pricing formula.

The proof is simply given by substituting the equation (10) into the equation (1). Substituting $\hat{\bar{F}}_{t,c}(m)$ into (12) we obtain a new method for pricing derivatives. The last term in (12) is the nonparametric correction of the pricing error due to the parametric pricing formula. The overall procedure is still nonparametric and can be combined with any parametric approach. We refer to this pricing approach as the Automatic Correction of Errors (ACE) approach.⁹ Such a bias reduction technique is effective in reducing pricing errors. This point will

⁹“Automatic” refers to the nonparametric fitting, which does not need to impose any form. This type of parametric guided approach has been used in the statistics literature; see, for example, Press and Tukey (1956) and Glad (1998).

be demonstrated by comparing the pricing performance of the ACE approach and the direct nonparametric method in the empirical applications in Section 3. As an example Figure 2 shows the estimates of the state price survivor function on December 29, 2004 using the call options analyzed in the previous sections and applying our ACE method and the direct nonparametric method. Visually the two estimates appear to be quite close and both methods seem to provide a good fit to the data. In contrast, Figure 3 shows that the pricing errors of the two methods behave very differently. The direct nonparametric approach does not perform well with an overall root mean square error (RMSE) of \$1.04, while our ACE approach has a RMSE of only \$0.21. The reason is that when pricing options via equation (1), the estimate of \bar{F} is multiplied by the current futures price, $F_{t,\tau}$, which was about \$1,200 on that day. Hence even small differences between the two estimates of the survivor function translate into substantial differences in the corresponding option prices. Figure 3 also presents the pricing errors of the ad hoc Black-Scholes model (6). This method does not perform well with a RMSE of \$1.10, mainly because the fitted volatilities $\hat{\sigma}^{\text{bs}}$ in equation (6) are not very accurate around the moneyness, $m \approx 1$. At-the-money options are most sensitive to changes in volatilities, having larger vega than in- and out-of-the-money options. Hence even small errors in the volatility estimation induce large errors in option prices. Figure 3 also shows the pricing performance of the semiparametric Black-Scholes model, to be introduced in Section 2.1, that is quite satisfactory with a RMSE of \$0.35. All the previous findings will be supported by the much more extensive empirical analysis in Section 3. Finally, as an example Figure 5 shows the estimate of the state price survivor function using the newly proposed method (10) on December 29, 2004, for four different time to maturities, $\tau = 24, 52, 80,$ and 171 days. For any given moneyness $m = X/F_{t,\tau}$, the estimate of $\bar{F}_t(m)$ gives the forward price of the digital call option paying one dollar if $F_{T,0} > X$ and zero otherwise. Such derivative contracts can be easily and accurately priced using our ACE method.

1.4 Adequacy of the pricing formula

Parametric pricing models are based on assumptions concerning the risk neutral dynamic of the underlying asset. Hence 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 distribution fit adequately the traded options? Statistically, this is a nonparametric hypothesis testing problem:

$$H_0 : \bar{F}_t(m) = \bar{F}_t(m; \theta) \quad \longleftrightarrow \quad H_1 : \bar{F}_t(m) \neq \bar{F}_t(m; \theta), \quad (13)$$

where $\bar{F}_t(m)$ is the true state price survivor function at time t for a given maturity T and $\bar{F}_t(m; \theta)$ is the state price survivor function derived from a stochastic parametric model. Note that the null hypothesis has a parametric form, while the alternative hypothesis is nonparametric. Hence the classical maximum likelihood ratio test needs to be properly extended to deal with such a general situation. One of such extensions is the generalized likelihood ratio (GLR) test proposed by Fan, Zhang, and Zhang (2001). For the current problem, the GLR test amounts to compare the residual sum of squares when fitting model (4) using the parametric and the nonparametric approaches. However, a direct application of the GLR statistic to the current problem is not ideal. Even when the null hypothesis is correct, the nonparametric fits incur biases; see Section 1.5

below. To improve the testing procedure, Fan and Yao (2003, Chapter 9) suggest to test whether or not the correction term $\bar{F}_{t,c}(m)$ is statistically significant away from zero. Under the null hypothesis that $\bar{F}_{t,c}$ is zero, $\bar{F}_t(m) = \bar{F}_t(m; \theta)$ and the residual sum of squares is

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

for a given sufficiently large interval $[a, b]$. Hence the survivor function is tested on the interval $[a, b]$ and the parameter, θ , characterizing $\bar{F}_t(m; \theta)$ is calibrated to fit the 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; see also Section 3.2. Similarly, under the alternative hypothesis of nonparametric model for the correction term, the residual sum of squares is

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

The GLR test statistics measures the inadequacy of the parametric fit and it is defined as

$$T_n = \frac{n_{a,b}}{2} \log(\text{RSS}_0/\text{RSS}_1), \quad (14)$$

where $n_{a,b} = \sum_{t=t_0-d}^{t_0+d} \sum_{i=1}^{N_t} I(a \leq \bar{m}_{t,i} \leq b)$ is the number of data points used in the fitting. Hence the larger the test statistics T_n is, the less adequate is the fit of the parametric model to 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 zero has to be rejected. To derive the asymptotic null distribution, we take $\lambda = 1$ for notational simplicity. Under the conditions listed in Appendix B, the asymptotic null distribution can be derived.¹⁰

Proposition 3 *Under the conditions in Appendix B, if the null hypothesis is true, then*

$$r_K T_n \stackrel{a}{\sim} \chi_{a_n}^2 \quad (15)$$

in the sense that

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

where, with $*$ denoting the convolution operator,

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

The constants r_K and s_K are computed in 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 demonstrates the adequacy of such an approximation. Fan and Yao (2003, Chapter 9) introduce a bootstrap method to better approximate the null distribution of the GLR test statistic. For simplicity and computational expediency, we use (15) to compute the P-value.

In the above formulation, the parametric model $\bar{F}_t(m; \theta)$ can be any survivor function. We took the log-normal survivor function (7) induced by the ad hoc Black-Scholes model as the parametric model in our

¹⁰For instance, similar assumptions are made by Ait-Sahalia and Lo (1998) and Gagliardini, Gouriou, and Renault (2005).

empirical analysis. The GLR test statistic (14) was applied to the S&P 500 index options traded from January 2, 2002 to December 31, 2004, as detailed in Section 3. We set the coefficients a and b equal to the 0.05 and 0.95 quantiles of the observed moneyness for each relevant maturity. We found that all test statistics have a P-value no larger than 0.001, providing stark evidence that the log-normal ad hoc Black-Scholes model $\bar{F}_{\text{LN}}(m; \vartheta)$ does not fit the options data well. As an example Figure 6 shows the residuals of the GLR test under the null and the alternative hypothesis using the call options observed on December 27–31, 2004 as in the previous sections. Especially around the moneyness, $m \approx 1$, the nonparametric correction term, $\bar{F}_{t,c}$, is quite effective in removing the bias in the residuals, $\tilde{Y}_{t,i}I(a \leq \bar{m}_{t,i} \leq b)$, and the null hypothesis that $\bar{F}_{t,c}$ is zero has to be rejected.

1.5 Statistical properties of ACE and bandwidth selection

We now show that the parametrically guided nonparametric fitting in Section 1.3 has smaller biases when the true state price survivor function is in the neighborhood of the parametric model. For this purpose, we assume that the parametric model on the survivor function is $\bar{F}(m; \theta)$ and that the least-square calibration method gives the θ that minimizes

$$\sum_{t=t_0-d}^{t_0+d} \sum_{i=1}^{N_t} (Y_{t,i} - \bar{F}(\bar{m}_{t,i}, \theta))^2. \quad (16)$$

Let $n = \sum_{t=t_0-d}^{t_0+d} N_t$ be the sample size. If the true survivor function during the time period $[t_0 - d, t_0 + d]$ is $\bar{F}_0(m)$,¹¹ then the nonlinear least-square (16) attempts to find θ_0 which minimizes

$$E(\bar{F}_0(m) - \bar{F}(m; \theta))^2. \quad (17)$$

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

Proposition 4 *Under the conditions given in Appendix C, we have*

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

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 $\sigma^2(m)$ is the conditional variance of $\varepsilon_{t,i}$ given $\bar{m}_{t,i} = m$.

The proof is given in Appendix C. From the above proposition, the bias of the parametric guided nonparametric estimator (i.e., an ACE method) has a leading term of order $\frac{1}{2} \bar{F}_c''(m) h^2 \int u^2 K(u) du$, while 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 the case when $\bar{F}(m, \theta_0)$ is close to the true state price survivor

¹¹Here, for ease of presentation, we assume that the true survivor function is the same from $t_0 - d$ to $t_0 + d$, or varies very slowly in short time periods. If this assumption is not valid, one needs to consider the date t_0 only, corresponding to $d = 0$. Similarly, we assume a generic parametric form $\bar{F}(m, \theta)$ also to simplify the presentation.

function. The advantages of a parametrically guided nonparametric regression over the direct nonparametric approach are also documented in Glad (1998) and Fan and Ullah (1999). For our particular application here, the curvature of $\bar{F}_0(m)$ is large when m is around one. 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 significantly reduced. Hence the ACE method performs better than the direct nonparametric approach.

Now we briefly discuss the issue of the bandwidth selection. Since the problem (9) is a standard nonparametric regression problem, a wealth of data-driven bandwidths can be employed; see Fan and Yao (2003). In particular, one can apply the pre-asymptotic 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 takes 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 $\{\bar{m}_{t,i}, i = 1, \dots, N_t; t = t_0 - d, \dots, t_0 + d\}$. The standard deviation accounts for the spreadness of the moneyness and the constant factor 0.3 is an empirical choice from trial-and-error.

2 Other pricing methods

The combination of model based pricing formulae and nonparametric learning is a new and powerful idea for pricing financial derivatives. To compare it more comprehensively with other ideas, we introduce two additional methods to be included in the empirical study. The first approach relaxes the quadratic form of the implied volatility in the ad hoc Black-Scholes model, allowing for more flexible curve fitting to volatility smiles. The second approach uses a parametric GARCH model to price options.

2.1 Semiparametric Black-Scholes model

As demonstrated in Figure 4, the parabola is not flexible enough to fit the implied volatility smile. One way to overcome this difficulty is to fit the implied volatility nonparametrically. This approach allows for a more flexible functional dependence of the implied volatility on moneyness, $\sigma^{\text{bs}}(m)$. Hence for each given date t_0 , we apply a local linear regression approach to estimate the implied volatility. This approach gives the flexibility of fitting a separate implied volatility curve for each maturity. As in the previous option pricing methods, we aggregate the data around the date t_0 to reduce the variability of the estimate and to enhance the continuity of the estimate across time.¹² For a given date t_0 , the local linear estimate $\hat{\sigma}^{\text{bs}}(m) = \hat{\beta}_0$ is given by the

¹²Aggregating the data around the date t_0 increases the sample size by an approximate factor of $2d + 1$. This will certainly reduce the variance of the resulting estimate. On the other hand, we only use data around the date t_0 . By the continuity of the state price density as a function of time, this will not introduce large estimation bias. Aggregation is really a time-domain smoothing resulting in a smoother estimated state price survivor function from one day to another. This is a desired property in practical implementation.

time-weighted local linear regression

$$\min_{\beta_0, \beta_1 \in \mathbb{R}^2} \sum_{t=t_0-d}^{t_0+d} \lambda^{|t_0-t|} \sum_{i=1}^{N_t} (\sigma_{t,i}^{\text{bs}} - \beta_0 - \beta_1 (m_{t,i} - m))^2 K_h(m_{t,i} - m), \quad (18)$$

where $m_{t,i}$ is the moneyness associated with the implied volatility, $\sigma_{t,i}^{\text{bs}}$, and $\hat{\beta}_0, \hat{\beta}_1$ are the resulting minimizers; see (5) for definitions of smoothing parameters λ , bandwidth h and kernel function K . As in the ad hoc Black-Scholes model (6), call option prices are computed by plugging $\hat{\sigma}^{\text{bs}}(m)$ in the Black-Scholes formula, $C_{t,i}^{\text{bs}}$, and setting $\sigma = \hat{\sigma}^{\text{bs}}(m)$. This pricing method is inspired by Ait-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 in the previous sections. The flexibility of the nonparametric fitting is evidenced.

2.2 GARCH option pricing model

The time varying volatility of financial asset returns is often described using GARCH models; see Engle (1982), and Bollerslev (1986) and for comprehensive surveys of GARCH models see Bollerslev, Engle, and Nelson (1994) and Ghysels, Harvey, and Renault (1996) among others. Several GARCH option pricing models have been proposed. In our empirical analysis we consider the Heston and Nandi (2000) model, as it is well-known in the financial literature. Heston and Nandi (2000) derive a closed-form option valuation formula when the spot asset S_t (including dividends) follows a GARCH model with Gaussian innovations. Under the risk neutral distribution

$$\begin{aligned} \log(S_t/S_{t-1}) &= r - h_t/2 + \sqrt{h_t} z_t \\ h_t &= \omega + \beta h_{t-1} + \alpha (z_{t-1} - \gamma \sqrt{h_{t-1}})^2, \end{aligned} \quad (19)$$

where z_t is a Gaussian innovation, h_t is the conditional variance of the log-return between $t-1$ and t and is known from the information available at time $t-1$. Heston and Nandi (2000) provide a detailed description of the model and here we only recall the main features. 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, respectively. When $\gamma > 0$, the model accounts for the leverage effect¹³ as a negative shock z_t raises the variance more than a positive shock z_t of the same absolute magnitude. At time $t = 0$, the call option C^{HN} with strike price X and maturity T is

$$\begin{aligned} C^{\text{HN}} &= e^{-rT} E^*[\max(S_T - X, 0)] \\ &= e^{-rT} \zeta(1) \left(\frac{1}{2} + \frac{1}{\pi} \int_0^\infty \Re \left[\frac{X^{-i\phi} \zeta(i\phi + 1)}{i\phi \zeta(1)} \right] d\phi \right) - e^{-rT} X \left(\frac{1}{2} + \frac{1}{\pi} \int_0^\infty \Re \left[\frac{X^{-i\phi} \zeta(i\phi)}{i\phi} \right] d\phi \right), \end{aligned} \quad (20)$$

¹³The name leverage effect was introduced by Black (1976) who suggested that a large negative return increases the financial and operating leverage, and rises equity return volatility; see also Christie (1982). Campbell and Hentschel (1992) suggested an alternative explanation based on the market risk premium and volatility feedback effects; see also the more recent discussion by Bekaert and Wu (2000). We shall use the name leverage effect as it is commonly used by researchers when referring to the asymmetric reaction of volatility to positive and negative return innovations.

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, $Y_T = \log(S_T)$,

$$\zeta(\phi) = E_t[e^{\phi Y_T}] = e^{\phi Y_t + A_t + B_t h_{t+1}}.$$

Heston and Nandi show that the moment generating function ζ has such a log-linear form under GARCH-type models; see also Christoffersen, Heston, and Jacobs (2006). The coefficients A and B are computed backward using the following equations and starting from the terminal condition $A_T = B_T = 0$,

$$\begin{aligned} A_t &= A_{t+1} + \phi r + B_{t+1} \omega - \frac{1}{2} \log(1 - 2\alpha B_{t+1}) \\ B_t &= \phi(\gamma - 1/2) - \frac{\gamma^2}{2} + \beta B_{t+1} + \frac{1/2(\phi - \gamma)^2}{1 - 2\alpha B_{t+1}}. \end{aligned}$$

The integrals in equation (20) are computed numerically via discretization. As integration domain we consider the interval $(0, 100)$ and we evaluate the integrand function on 5,000 equally spaced mid-points. Then the integrals are computed by averaging the function values over such an interval.¹⁴ As the moment generating function, ζ , does not depend on the strike X , for each given maturity T the coefficients A and B need to be computed only once. Hence this numerical procedure allows to evaluate a whole cross section of option prices in a few seconds, largely reducing the calibration time. Furthermore, given the past underlying returns the current variance, h_{t+1} , is known at time t . Hence the pricing formula (20) does not require the calibration or the separate estimation of h_{t+1} , simplifying the calibration procedure. This is an important advantage of GARCH pricing models over other stochastic volatility models, such as the Heston (1993) and the Bakshi, Cao, and Chen (1997) models.

For each Wednesday in our sample, we calibrate the GARCH model (19) to the cross section of call options by minimizing the squared pricing errors

$$\{\omega_t, \beta_t, \gamma_t, \alpha_t\} = \arg \min_{\omega, \beta, \gamma, \alpha} \sum_{j=1}^{M_t} (C_{t,j}^{\text{HN}}(\omega, \beta, \gamma, \alpha) - C_{t,j})^2,$$

where $C_{t,j}$ denotes the market option price and M_t is the number of options traded at time t . An important advantage of this approach is that it allows to infer the GARCH parameters $\{\omega, \beta, \gamma, \alpha\}$ using all options traded at time t with different maturities, and hence aggregating all the cross sectional information. The disadvantage is that the calibration procedure is quite computationally demanding. The highly nonlinear pricing formula (20) has to be re-evaluated at each step of the optimization procedure and the danger of finding local minimizers instead of the global one is a concern.¹⁵ The justification of the GARCH model in the risk neutral world is another mathematical challenge.¹⁶ One advantage of our ACE approach proposed

¹⁴The particular integration domain and subinterval length are selected to match option prices on a few selected days computed using the Romberg's numerical integration method over the interval $(10^{-6}, 200)$ with tolerance 10^{-4} . Some differences between the two approaches are observed only for a few deepest out-of-the-money call option prices.

¹⁵In our empirical application to solve the calibration problem we use the Nelder-Mead simplex direct search method as implemented in the Matlab function `fminsearch`.

¹⁶Duan (1995) discusses the preference assumptions that induce the risk neutral distribution in a GARCH model with Gaussian innovations; see for instance Stutzer (1996) for related work on the preference assumptions characterizing the risk neutral density.

in Section 1.3 is that it can be combined with any approach to enhance the accuracy of a pricing formula. For instance regarding $C^{\text{HN}}(\hat{\omega}, \hat{\beta}, \hat{\gamma}, \hat{\alpha})$ as a preliminary estimate, its pricing errors can be nonparametrically learned and corrected, but we do not pursue this direction in the current paper.

In Section 3 we compare the pricing and hedging performances of the previous option pricing methods. The GARCH model relies on a computationally intensive procedure, but it is calibrated only once a week to the cross section of options. All the other methods are much less computationally intensive than the GARCH model and are separately fitted to options with different maturities.

3 Empirical analysis

3.1 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, downloaded from OptionMetrics. The average of bid and ask prices are taken as option prices and options with time to maturity less than 20 days or more than 240 days,¹⁷ implied volatility larger than 70%, or prices less than 1/8 are discarded, which yields a sample of 101,036 observations.

The market for SPX options is one of the most active index options market in the world. 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 and have no wild card features. Furthermore, SPX options can be hedged using the active market on the S&P 500 futures. Consequently, SPX options have been the focus of many empirical investigations, including Ait-Sahalia and Lo (1998), Chernov and Ghysels (2000), Heston and Nandi (2000), Carr, Geman, Madan, and Yor (2003), and Barone-Adesi, Engle, and Mancini (2006). The term structure of default-free interest rates is also downloaded from OptionMetrics and the riskless interest rate for each given 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.

The raw data presents three challenges. First, in-the-money options are not actively traded compared to at-the-money and out-of-the-money options. For instance, the daily volume for out-of-the-money puts is usually several times as large as the volume for in-the-money puts. This phenomenon started after the October '87 crash and reflects the strong demand by portfolio managers for protective puts, inducing the well-known implied volatility smile; see for instance Figure 4. Second, it is difficult to observe the underlying index price exactly when option prices are recorded. Even slight temporal mismatches between option and index price recordings can induce pricing biases; see, for instance, Fleming, Ostdiek, and Whaley (1996) and Whaley (1993, Appendix). Third, the S&P 500 index pays dividends and the future rate of dividend is difficult to

¹⁷The lower bound of option maturities is mainly motivated by our estimation strategy, (aggregating options over one week), and by the out-of-sample exercise in Section 3.4, (predicting options one week ahead). For short time to maturity, the option price sensitively depends on the change of the time to maturity. Hence the continuity of the option price with respect to the time to maturity is harder to explore. The upper bound is mainly motivated by liquidity arguments.

determine. Daily dividend payments for the index are available from the S&P 500 Information Bulletin, but by nature these data are backward-looking and it is hard to assume that actual dividends recorded ex post correctly reflect expected future dividends; see Harvey and Whaley (1991). We address these three problems following the procedure suggested by Ait-Sahalia and Lo (1998). Because option prices are recorded at the same time on each day, only one temporally matched index price per day is required to estimate a pricing model. To circumvent the problem of non observable dividend yield, $\delta_{t,\tau}$, the futures price $F_{t,\tau}$ for each maturity τ is derived. $F_{t,\tau}$ and S_t are linked by the spot-futures parity

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

The futures price $F_{t,\tau}$ is computed via the put-call parity relation which holds because of the absence of arbitrage opportunity, independently of any option pricing model

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

where P_t denotes the put price. The futures price $F_{t,\tau}$ is inferred using liquid calls and puts closest to at-the-money with same strike price X and time to maturity τ . The procedure is repeated for all dates and time to maturities τ obtaining the implied futures prices from put-call parity for each maturity. Given the derived futures price $F_{t,\tau}$, the prices of illiquid in-the-money call options are replaced by the corresponding prices implied by the put-call parity, $P_t + F_{t,\tau} e^{-r_{t,\tau}\tau} - X e^{-r_{t,\tau}\tau}$, where the put price is out-of-the-money and therefore liquid. After this procedure the information in liquid out-of-the-money put prices translates into implied in-the-money call prices via the put-call parity. Hence put prices may be discarded without any loss of reliable information. Moreover, some empirical studies investigate model pricing performances separately using calls and puts, and the empirical findings based on calls and puts are rather similar; see, for instance, Bakshi, Cao, and Chen (1997) and Dumas, Fleming, and Whaley (1998).

We divide the 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 $m < 0.8$; in-the-money (ITM) if $0.8 \leq m < 0.94$; at-the-money (ATM) if $0.94 \leq m < 1.04$; out-of-the-money (OTM) if $1.04 \leq m < 1.2$; and deep out-of-the-money (DOTM) if $m \geq 1.2$. An option contract can be classified, by the time to maturity, as short maturity (< 60 days); medium maturity (60–160 days); and long maturity (> 160 days). Table 1 describes the 101,036 call option prices and the corresponding implied volatilities used in the empirical analysis after applying the previous filters and replacing illiquid in-the-money call options. 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, respectively, 20, 22, and 21 percent of the total sample. Short and long maturity options account for, respectively, 37 and 20 percent of the total sample. The table also shows the volatility smile and the corresponding term structure. For each given set of maturities, the smile across moneyness is evident. Furthermore, the smile tends to become flatter when time to maturity increases.

3.2 Implementing the option pricing models

For each Wednesday from January 2, 2002 to December 31, 2004, we calibrate the GARCH pricing model (19) to the cross section of options. Aggregating data over each week, for each Wednesday and each maturity we fit our ACE approach (12), the direct nonparametric model (5), the semiparametric Black-Scholes model (18), and the ad hoc Black-Scholes model (6). Then we compute the prices of the European options available on each Wednesday obtaining 16,521 option price estimates for each model.

In the nonparametric regressions, we use the local linear approximation and the Epanechnikov kernel with bandwidth $h = 0.3s$, where s is the sample standard deviation of the moneyness; see the end of Section 1.4 for a discussion on the bandwidth selection. Indeed, we also experimented other similar bandwidth values and the overall results for the nonparametric and semiparametric methods were largely the same. We set $\lambda = 10^{-30/365} \approx 0.83$, and we measure the distance $|t - t_0|$ in calendar days. This choice assigns the weights 0.68, 0.83, 1, 0.83, 0.68 to options traded on the different week days, from Monday to Friday. We use the same time-weights for our ACE approach (12), the direct nonparametric model (5), the semiparametric Black-Scholes model (18) and the ad hoc Black-Scholes model (6). We also experimented other values for λ and magnitude within reason we obtained very similar results for all the pricing models. The correction term $\bar{F}_{t,c}$ in the ACE method (12) is estimated from the 0.05 to the 0.95 quantiles of the observed moneyness and beyond this interval $\bar{F}_{t,c}$ is set to zero. This procedure ensures that the estimation of the correction term is based on a sufficiently large number of observations. We applied the same procedure in the GLR test analysis for the adequacy of the pricing formula in Section 1.4.

When implementing the GARCH pricing formula (20), the dividends paid by the stocks in the S&P 500 index have to be taken into account.¹⁸ For each maturity, we compute the dividend yield using spot-futures and put-call parities based on liquid closest at-the-money options, and then we subtract it from the current index level. As in Christoffersen, Heston, and Jacobs (2006), we set the riskless rate $r = 0.05/365$ in the GARCH pricing formula (20) and at each step of the calibration procedure, the conditional variance, h_t , is initialized at the unconditional variance level, $(\omega + \alpha)/(1 - \beta - \alpha\gamma^2)$, and then updated using the GARCH dynamic (19). The iteration is started 250 trading days before the first option date to allow for the model to find the right conditional variance.

3.3 In-sample model comparisons

We first investigate the pricing performance of each option pricing model using in-sample data. 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 according to the root mean square and absolute pricing errors measured both in dollar and in relative terms. Tables 3 and 4 disaggregate the dollar and the relative pricing errors across the five moneyness and three maturities categories. The ACE method has the lowest RMSE for most comparisons, which explains the superior overall pricing performance in Table 2. It

¹⁸All the other pricing methods are implemented using futures prices, which already embed dividend yields.

has some difficulties in pricing deep in-the-money options and the reason is that to price options with low moneyness, such as $m \approx 0.4$, the survivor function has to be integrated for almost all the moneyness domain, accumulating all the estimation errors in $\bar{F}_t(m)$. However, deep in-the-money options are quite “expensive” and not actively traded. Interestingly, the ACE method performs well in pricing more actively traded at-the-money and out-of-the-money options.

As shown in Tables 2, 3 and 4, the second best method is the semiparametric Black-Scholes model (18) (label Semip-BS), which performs particularly well for deep in-the-money options. Although the ad hoc Black-Scholes model (6) (label Ad Hoc BS) is estimated for each maturity, it does not perform well mainly because the volatility smiles for large moneyness ranges are not well approximated by parabolic functions. The semiparametric Black-Scholes model is designed to circumvent this problem and always outperforms the ad hoc Black-Scholes model. Table 10 shows summary statistics of the parabola coefficients and confirms that the volatility smile is a stable and persistent characteristic of implied volatilities.

The direct nonparametric model (5) (label NP) does not perform well, as anticipated, because a constant bandwidth is not adequate to estimate the survivor function over the whole domain. Moreover, it does not take into the account the prior knowledge on the shape of the state price survivor function. The proposed bias reduction technique incorporated in the ACE method is effective in reducing pricing errors. The tables provide stark evidence for the power of our idea of combining the explanatory power of model based pricing formulae and nonparametric learning to correct pricing errors. None of the four methods alone can be placed ahead of the ACE approach. The ACE method performs spectacularly well for call options when the moneyness is not too small and the time to maturity is not too short. One can reliably use this method for pricing derivatives in these categories.

The GARCH pricing model (19) (label GARCH) does not perform well, mainly because it lacks of flexibility to separately fit options with different maturities when compared for instance to the ACE method. However, the GARCH model tends to outperform the ad hoc Black-Scholes model for options around the moneyness and short-term maturities. Table 11 shows summary statistics of the calibrated GARCH parameters. As in previous studies (for instance Heston and Nandi (2000) and Barone-Adesi, Engle, and Mancini (2006)), the parameter γ is largely positive confirming the leverage effect, i.e. the stronger impact of negative shocks than positive shocks in raising the variance level. Clearly, the GARCH parameters change over time. In contrast, Table 12 shows the corresponding estimates of the long run risk neutral volatility and the persistency of the variance process, $\beta + \alpha\gamma^2$. Both estimates are much more stable over time than individual parameter estimates and in line with other empirical studies.

Figure 7 shows graphically the in-sample absolute dollar and relative pricing errors across moneyness and maturities categories for the different pricing models.¹⁹ Clearly, the two methods that we proposed, the ACE and the semiparametric Black-Scholes methods outperform all the other pricing models. Furthermore, for most moneyness and maturity categories the ACE method outperforms the semiparametric Black-Scholes method.

¹⁹Given the relatively poor pricing performance of the direct nonparametric model the corresponding plot is omitted in the in-sample comparison as well as in the subsequent out-of-sample and hedging comparisons.

In the previous analysis the nonparametric and semiparametric methods are estimated on each Wednesday using also future information, namely options data on Thursday and Friday. This approach allows to exploit the time continuity of the pricing function, as for instance in Ait-Sahalia and Lo (1998). However, the previous methods can also be implemented using only current and past data. As a robustness check we repeated the previous in-sample analysis (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 new results (not reported here) largely confirm the current findings and are available from the authors upon request. Indeed for almost all the moneyness and maturity categories—for all the in-sample, out-of-sample and hedging analysis—the performance ranking of the different models coincides with the current ranking. In the following section we undertake the out-of-sample analysis and each pricing model is evaluated without borrowing future information. As the direct nonparametric model is largely outperformed by the ACE method and also by the semiparametric and the ad hoc Black-Scholes models, its out-of-sample and hedging analysis will be omitted.

3.4 Out-of-sample model comparisons

Out-of-sample forecasting of option prices is an interesting challenge for any pricing method. It tests not only the goodness-of-fit of the pricing formula, but also whether or not nonparametric methods overfit the option prices in the in-sample learning period. On each Wednesday and for each maturity, in-sample model estimates are used to price the same options one week later using futures prices, time to maturities and interest rates relevant on the next Wednesday.²⁰ The state price survivor functions estimated on the current Wednesday will be used to forecast option prices traded on the next Wednesday. Some small modifications are needed for the ACE method. The model based pricing, i.e. the log-normal model in the current implementation, needs to take into account the change of the time to maturity. If the Black-Scholes formula was used as the model based price, then the coefficient accounting for the change of the time to maturity was $\vartheta_t = \sqrt{T-t}$. To forecast the effects due to the change of the time to maturity, we fit the following simple regression model

$$\hat{\vartheta}_t = \kappa\sqrt{T-t} + \text{error}_t, \quad \text{for } t = t_0 - d, \dots, t_0 + d, \quad (21)$$

where $\hat{\vartheta}_t$'s are given by (8). Equation (21) links time to maturities and $\hat{\vartheta}_t$'s in the state price survivor function, $\bar{F}_{LN}(m; \hat{\vartheta}_t)$. Using in-sample data observed over one week or five trading days, 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 that reflects the shorter time to maturity at the future date t_1 is $\hat{\vartheta}_{t_1} = \hat{\kappa}\sqrt{T-t_1}$. In our application $t_1 = t_0 + 7$ days.²¹

²⁰The one week ahead forecast horizon is also adopted by Dumas, Fleming, and Whaley (1998) and Heston and Nandi (2000), among others and for comparative purposes we adopt it here as well.

²¹The coefficient $\hat{\vartheta}_{t_1}$ could be computed solving the minimization problem (8) using options data on time t_1 , but this would invalidate the out-of-sample analysis.

The out-of-sample pricing performances are summarized in Table 5 and disaggregated by moneyness and maturity in Tables 6 and 7. Interestingly, the results share the same pattern as the in-sample pricing results. In particular, it is evidenced that the ACE method outperforms all the other pricing methods in terms of dollar and relative pricing errors, sometimes even to a larger extent. These results demonstrate that the automatic bias correction in the ACE is very effective. Without this part, ACE is essentially the same as the ad hoc Black-Scholes method, which does not perform well. As explained before, the ACE method is particularly accurate in pricing options when the moneyness is not too small because less estimation errors in $\bar{F}(m)$ are accumulated in the integral (1). These tables also show that the second best method is the semiparametric Black-Scholes model, although the relative pricing errors increase when the time to maturities decrease. It is known that short term and especially out-of-the-money option prices are hard to predict due to high volatilities. Figure 8 shows the out-of-sample absolute dollar and relative pricing errors for all the pricing models. It is clear that in dollar terms it is harder to predict at-the-money options than out-of-the-money ones. In relative terms, call options with large moneyness are hard to price.

When predicting option prices in the GARCH setting, the conditional variance, h_t , is updated using the current calibrated risk neutral parameters and the actual S&P 500 daily log-returns from t_0 to $t_0 + 7$ days. Tables 5, 6 and 7 show that, also in the out-of-sample exercise, the GARCH model has larger prediction errors than the ACE and the semiparametric Black-Scholes methods. These findings confirm the power of nonparametric approaches. For at-the-money options with short-term maturities the GARCH model continues to outperform the ad hoc Black-Scholes model.

3.5 Hedging results

An important motivation for developing option pricing models is to provide better management of the risk involved in options. Setting hedge ratios based on accurate and reliable valuation model should induce an improvement in the hedging performance. Following Dumas, Fleming, and Whaley (1998), we evaluate the performance of a hedge portfolio formed on day t and liquidated one week later on day $t + 7$. It is known that the return on such a discretely adjusted option hedge portfolio has three components: (i) the risk-free return on investment, (ii) the return from the discrete adjustment of the hedge, and (iii) the return from the difference between the change in the actual option price and the change in the theoretical option price over the week horizon; see Galai (1983) and Dumas, Fleming, and Whaley (1998). As in our analysis we use forward option prices, the risk-free return component of the hedge portfolio is zero. Furthermore, as our focus is on model performance and not on the issues raised by discrete time recalibration, we assume that the hedge portfolio is continuously rebalanced through time. Hence the hedging error is defined as

$$\epsilon_t = \Delta C_{\text{actual},t} - \Delta C_{\text{model},t}, \quad (22)$$

where $\Delta C_{\text{actual},t}$ is the observed change in the market option price from date t to date $t + 7$ and $\Delta C_{\text{model},t}$ is the change in the model theoretical price. The proof of equation (22) is provided by Dumas, Fleming, and Whaley (1998, Section VI) and for completeness we briefly recall it here as well. The hedging error resulting

from the continuous time recalibration using the hedge ratio, ρ , is

$$\Delta C_{\text{actual},t} - \int_t^{t+\tau} \rho(S_u, u) dS_u. \quad (23)$$

If the valuation model gives the correct hedge ratio, ρ , to continuously rebalance the hedge, the two terms in equation (23) would be equal with probability one and the integral term must be equal to $\Delta C_{\text{model},t}$. In other words, if the valuation model is not correct the hedging error is equal to the time increments in the valuation error.

Table 8 summarizes the hedging errors of the four pricing models across the different years from 2002 to 2004 and Table 9 disaggregates the hedging results across moneyness and maturity. The hedging results largely confirms the previous in- and out-of-sample pricing results. Our ACE approach tends to outperform all the other methods sometimes by a large extent, but it has some difficulties in hedging very short maturity, deep in-the-money options. The overall second best method is the semiparametric Black-Scholes model, while the GARCH model tends to be dominated by the ad hoc Black-Scholes model. Figure 9 shows the absolute hedging errors of the pricing models and visually confirms the previous results.

4 Conclusions

We propose a new nonparametric method for estimating state price distributions and pricing financial derivatives. This method is called the Automatic Correction of Errors (ACE) in a pricing formula. The ACE approach is based on a physical model guided nonparametric estimate of the state price distribution. Given any parametric pricing model, the corresponding pricing errors are nonparametrically learned and corrected thorough the estimate of the state price distribution, improving the model pricing performance. The ACE method is easy to implement and can be combined with any model based pricing formula to correct the systematic biases of pricing errors and enhance the predictive power. We also propose a semiparametric Black-Scholes method for option pricing, ameliorating an approach of 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. Finally, the ACE approach could be applied also in other contexts. For instance an interesting application could be in credit risk modeling where the accurate estimation of the default survivor distribution is essential in the pricing of credit risk sensitive contingent claims. This accuracy could be achieved by applying the ACE approach.

		Maturity					
		Less than 60		60 to 160		More than 160	
		Mean	Std.	Mean	Std.	Mean	Std.
DITM	Call price \$	323.22	89.16	342.00	108.12	354.06	117.30
	$\sigma^{\text{bs}}\%$	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
	$\sigma^{\text{bs}}\%$	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
	$\sigma^{\text{bs}}\%$	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
	$\sigma^{\text{bs}}\%$	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
	$\sigma^{\text{bs}}\%$	37.38	11.55	25.91	8.40	20.32	4.72
	Observations	4,561		8,775		4,783	

Table 1: Database description. The table shows mean, standard deviation (Std.) 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 in-the-money options as described in the main text. σ^{bs} is the Black-Scholes implied volatility. DITM is deep in-the-money options with moneyness less than 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 larger than 1.2. Moneyness is defined as strike price divided by the futures price. Maturity is measured in calendar days.

Panel A: Aggregated valuation errors across all years

	Bias	RMSE	MADE	Min	Max	Err>0%	Bias%	RMSE%	MADE%
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

	Bias	RMSE	MADE	Min	Max	Err>0%	Bias%	RMSE%	MADE%
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

Table 2: In-sample pricing errors. Bias, root mean square error (RMSE) and mean absolute error (MADE) of the dollar pricing error, (model price – market price), and of the percentage relative pricing error, $100 \times (\text{model price} - \text{market price}) / \text{market price}$; 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.

		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

Table 3: In-sample dollar pricing errors disaggregated by moneyness and maturity. Bias and root mean square error (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.

		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

Table 4: In-sample percentage relative pricing errors disaggregated by moneyness and maturity. Bias and root mean square error (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.

Panel A: Aggregated valuation errors across all years

	Bias	RMSE	MADE	Min	Max	Err>0%	Bias%	RMSE%	MADE%
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

	Bias	RMSE	MADE	Min	Max	Err>0%	Bias%	RMSE%	MADE%
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

Table 5: Out-of-sample pricing errors. Bias, root mean square error (RMSE) and mean absolute deviation error (MADE) of the dollar pricing error, (model price – market price), and of the percentage relative pricing error, $100 \times (\text{model price} - \text{market price}) / \text{market price}$; 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.

		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

Table 6: Out-of-sample dollar pricing errors disaggregated by moneyness and maturity. Bias and root mean square error (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.

		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

Table 7: Out-of-sample percentage relative pricing errors disaggregated by moneyness and maturity. Bias and root mean square error (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.

Panel A: Aggregated hedging errors across all years

	Mean	RMSE	MADE	Min	Max	Err>0%
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

	Mean	RMSE	MADE	Min	Max	Err>0%
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

Table 8: Hedging error. Mean, root mean square error (RMSE) and mean absolute deviation error (MADE) of the dollar hedging error in equation (22); 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.

		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

Table 9: Hedging errors disaggregated by moneyness and maturity. Mean and root mean square error (RMSE) of the dollar hedging error in equation (22) 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.

Year	a_0		a_1		a_2	
	Mean	St. Dev.	Mean	St. Dev.	Mean	St. Dev.
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

Table 10: Ad hoc Black-Scholes model (6) calibrated on each Wednesday from January 2, 2002 to December 31, 2004, using call option prices.

Year	$\omega \times 10^{14}$		β		γ		$\alpha \times 10^6$	
	Mean	St. Dev.	Mean	St. Dev.	Mean	St. Dev.	Mean	St. Dev.
2002	2.53	7.60	0.67	0.14	246.15	73.27	5.79	3.66
2003	1.07	3.96	0.66	0.22	284.69	93.57	5.48	6.89
2004	0.82	3.09	0.64	0.14	563.63	361.61	2.41	2.71

Table 11: Heston and Nandi GARCH model (19) calibrated on each Wednesday from January 2, 2002 to December 31, 2004, using call option prices; parameters on a daily base.

Year	$E[h^{1/2}]$		$\beta + \alpha\gamma^2$	
	Mean	St. Dev.	Mean	St. Dev.
2002	0.23	0.04	0.96	0.03
2003	0.22	0.03	0.95	0.07
2004	0.19	0.05	0.96	0.06

Table 12: Long run risk neutral volatility on an annual base, $E[h_t^{1/2}] = \sqrt{365(\omega + \alpha)/(1 - \beta - \alpha\gamma^2)}$, and persistency, $\beta + \alpha\gamma^2$, of the Heston and Nandi GARCH model (19) calibrated on each Wednesday from January 2, 2002 to December 31, 2004, using call option prices. Recall that $\beta + \alpha\gamma^2 = 1$ implies an integrated variance process h_t .

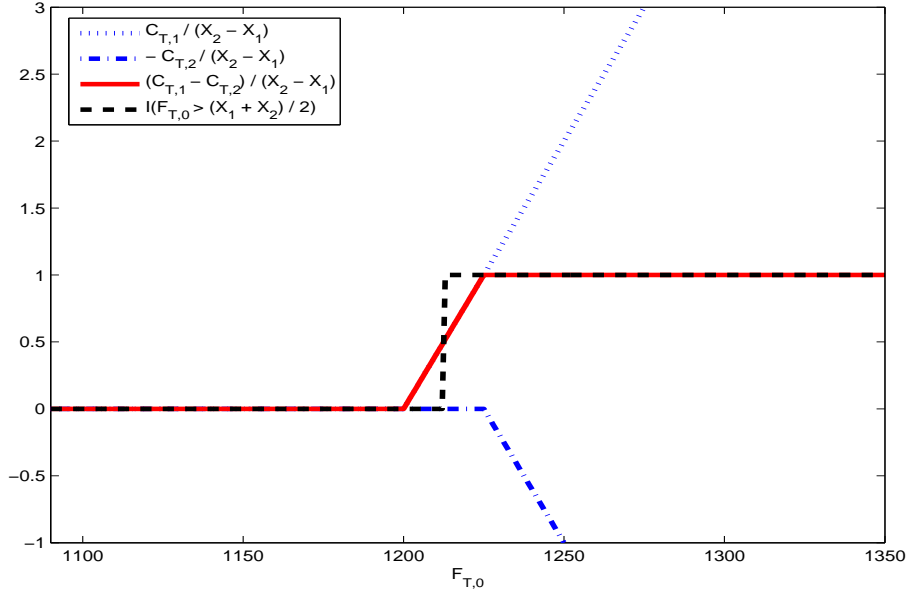


Figure 1: The pay-off function (solid line) with long 0.04 shares on the call option with strike price $X_1 = \$1,200$ and short 0.04 shares on the call option with strike price $X_2 = \$1,225$ at the time of maturity is basically an indicator function (dashed line). The expected value of such a portfolio is approximately the same as the state price survivor function (one minus the state price distribution) evaluated at $x = (X_1 + X_2)/2$. The dotted and dash-dotted lines are the pay-off functions of the long and short call options.

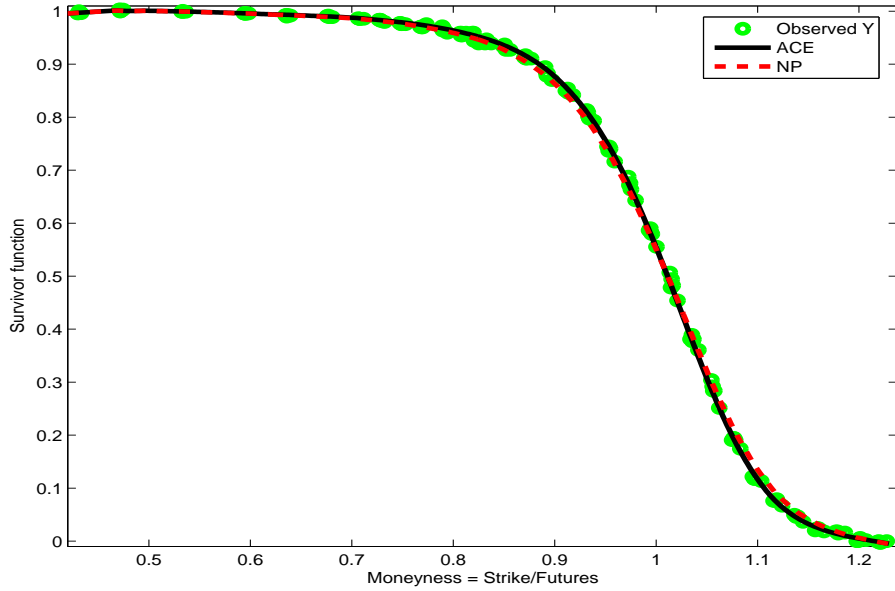


Figure 2: Scatter plot of $Y_{t,i} = e^{r\tau}(C_{t,i} - C_{t,i+1})/(X_{t,i+1} - X_{t,i})$ versus the moneyness, $\bar{m}_{t,i}$, for the call option prices observed on December 27–31, 2004 with maturities 169–173 days. The survivor function is estimated by using the direct nonparametric approach (5) (NP) and the newly proposed ACE approach (10).

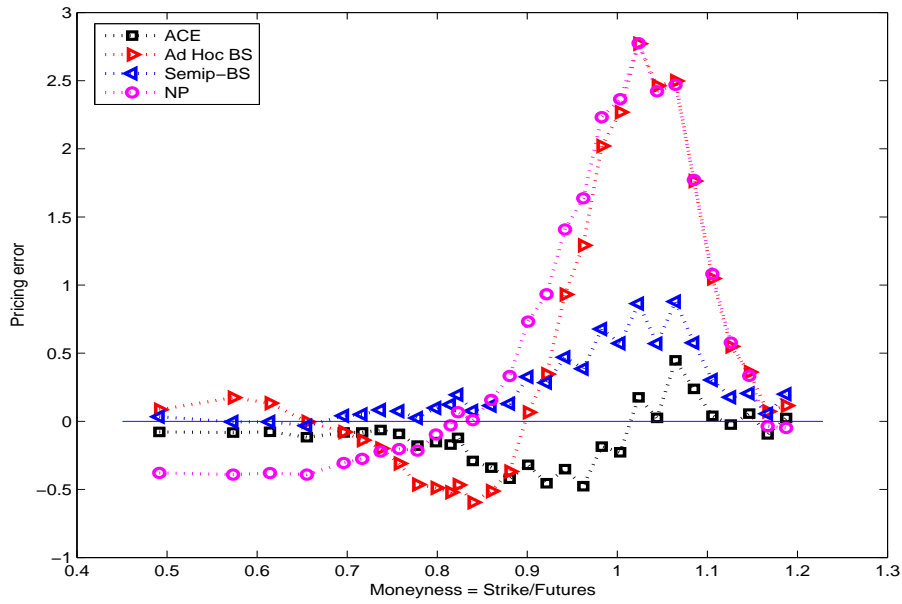


Figure 3: Pricing performance of the the direct nonparametric approach (NP), the Ad Hoc Black-Scholes model (Ad Hoc BS), the ACE approach, the semiparametric Black-Scholes (semip-BS) formula. The picture shows the corresponding dollar pricing error, (model price – market price), on December 29, 2004.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

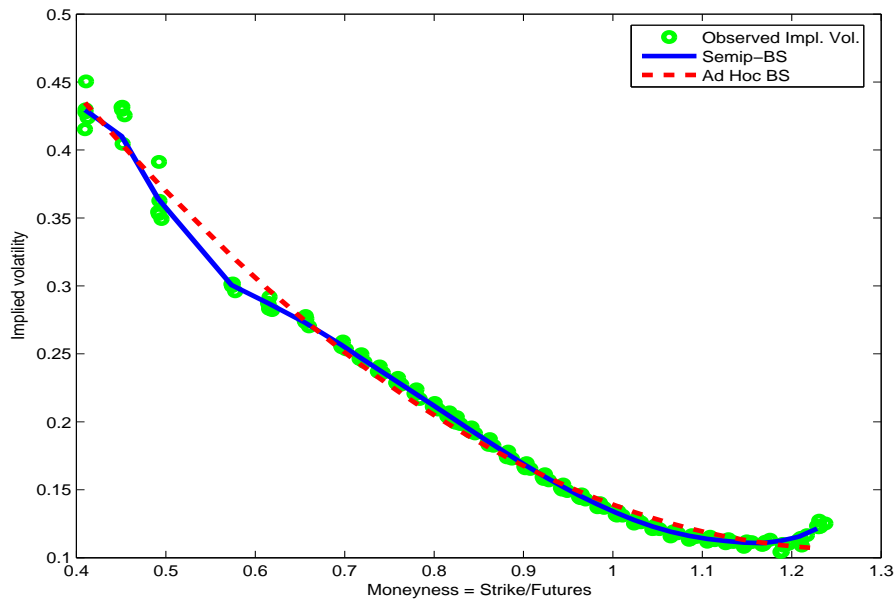


Figure 4: Implied volatilities observed on December 27–31, 2004 for call option prices with maturities 169–173 days. The figure shows the fitted parabola, $\sigma = a_0 + a_1m + a_2m^2$, (dashed line), and the local linear estimation of the function $\sigma(m)$, (solid line), where m denotes the moneyness.

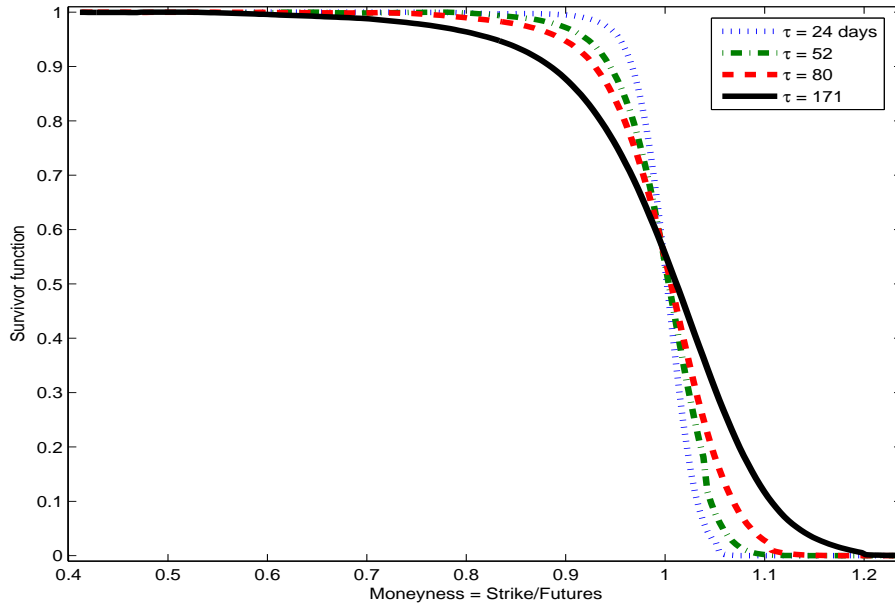


Figure 5: State price survivor functions estimated using the newly proposed ACE approach (10) on December 29, 2004 for four different maturities.

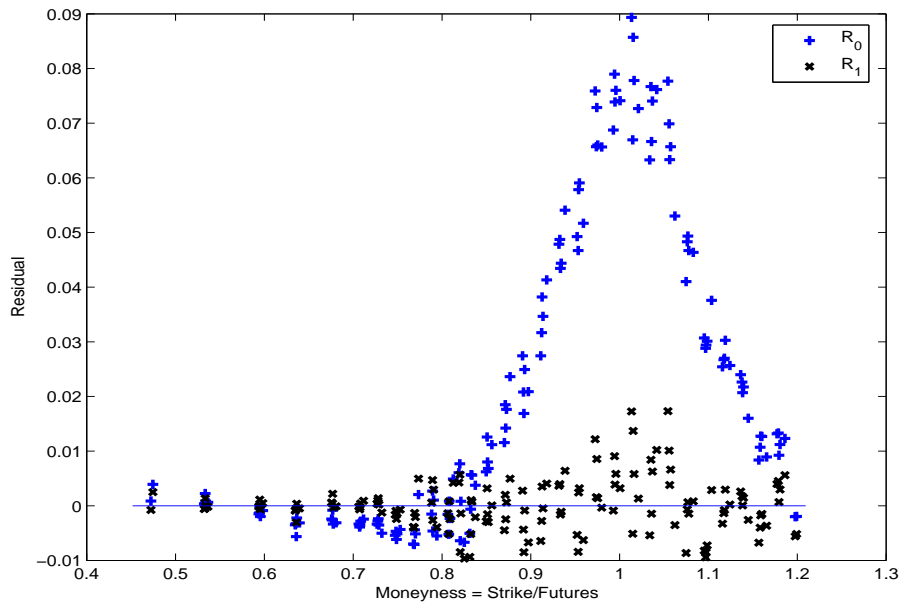


Figure 6: Residuals of the GLR test in equation (15) using option prices observed on December 27–31, 2004. R_0 's are the residuals under the parametric null hypothesis $H_0 : \bar{F}_{t,c} = 0$, i.e. $R_0 = \tilde{Y}_{t,i} I(a \leq \bar{m}_{t,i} \leq b)$, and R_1 's are the residuals under the alternative nonparametric hypothesis $H_1 : \bar{F}_{t,c} \neq 0$, i.e. $R_1 = (\tilde{Y}_{t,i} - \bar{F}_{t,c}(\bar{m}_{t,i})) I(a \leq \bar{m}_{t,i} \leq b)$. a and b are the 0.05 and 0.95 quantile of the observed moneyness range.

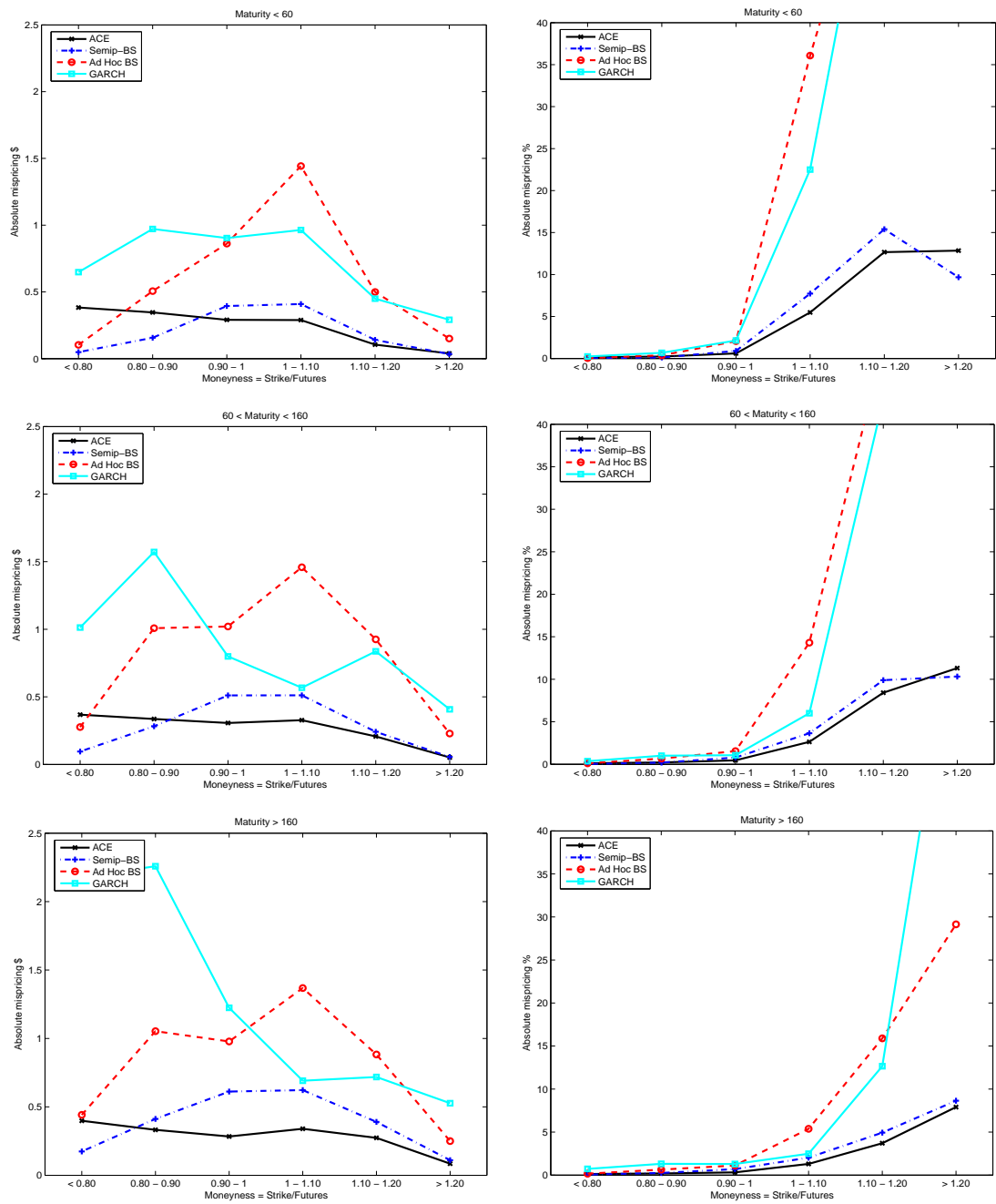


Figure 7: In-sample absolute mispricing in dollars, i.e. $|\text{model price} - \text{market price}|$, (left plots), and in percentage, i.e. $100 \times |\text{model price} - \text{market price}| / \text{market price}$, (right plots), for the different pricing models, averaged across the Wednesdays from January 2, 2002 to December 31, 2004 for the SPX call options.

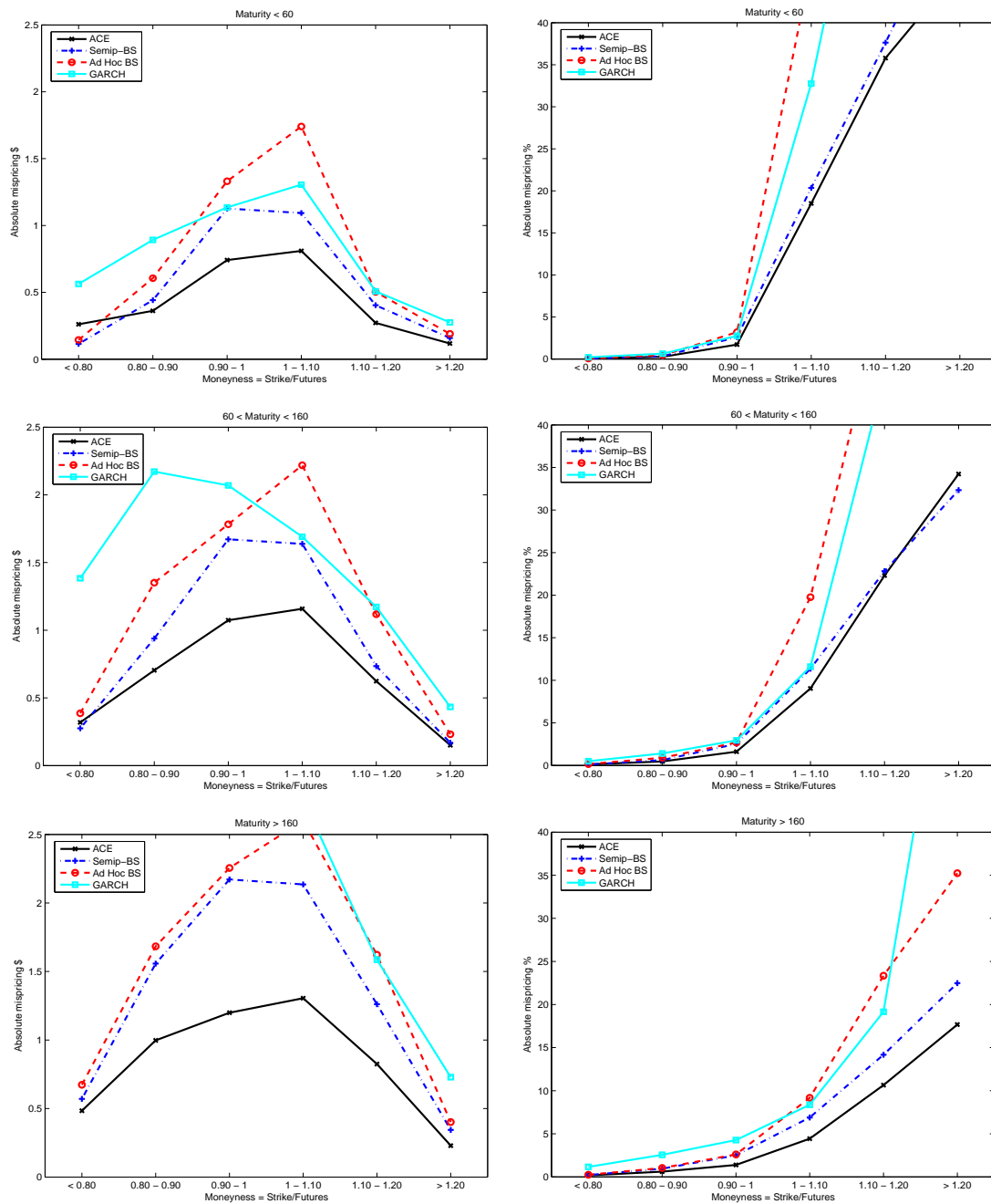


Figure 8: Out-of-sample absolute mispricing in dollars, i.e. $|\text{model price} - \text{market price}|$, (left plots), and in percentage, i.e. $100 \times |\text{model price} - \text{market price}| / \text{market price}$, (right plots), for the different pricing models, averaged across the Wednesdays from January 2, 2002 to December 31, 2004 for the SPX call options.

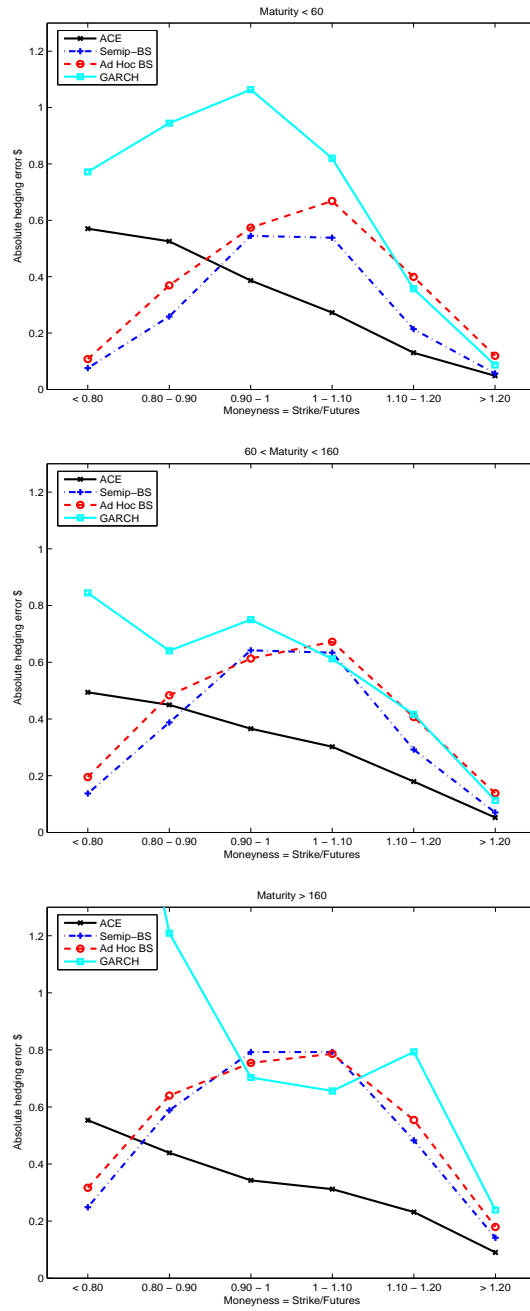


Figure 9: Absolute hedging error in dollars for the different pricing models, averaged across the Wednesdays from January 2, 2002 to December 31, 2004 for the SPX call options.

A Proof of Proposition 1

By (1), the left hand side of (3) 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. As $\bar{F} = 1 - F$, $\bar{F}' = -f$ and $\bar{F}'' = -f'$, by 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 zero, which is the advantage of using the mid-point, \bar{m} , and the second integral is bounded by

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

This completes the derivation of Proposition 1.

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 i.i.d. random variables, satisfying model (4) with i.i.d. homoscedastic random noise, $\varepsilon_{t,i}$. We assume further that

(A1) The marginal density of moneyness $\{m_{t,i}\}$ is bounded away from zero 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. Further the functions $t^3 K(t)$ and $t^3 K'(t)$ are bounded and $\int t^4 K(t) dt < \infty$.

(A5) The bandwidth h satisfies $h \rightarrow 0$ and $nh^{3/2} \rightarrow \infty$.

(A6) The function $F(x; \theta)$ has a continuous derivative with respect to θ and the calibrated $\hat{\theta}$ is root-n consistent.

Under the above 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. $n = \sum_{t=t_0-d}^{t_0+d} N_t$. Now, letting $T'_n = \frac{n}{2} \log(\text{RSS}_0/\text{RSS}_1)$, we have

$$T_n = n^{-1} n_{a,b} T'_n = P(m \in [a, b]) T'_n + O_P(n^{-1/2} T'_n). \quad (24)$$

We can apply the result of Fan, Zhang, and Zhang (2001) to T'_n , noting that the null hypothesis using our method is $\bar{F}_{t,c} = 0$. In particular, according to the Remark 4.2 of Fan, Zhang, and Zhang (2001), we have $r'_K T'_n \stackrel{a}{\sim} \chi_{a_n}^2$, where

$$r'_K = r_K P(m \in [a, b]), \quad a_n = s_K(b-a)/h.$$

JQ: Should the last a_n be a'_n ? Or $r'_K T'_n \stackrel{a}{\sim} \chi_{a_n}^2$, without the ' in a_n ? In Proposition 3 we have $a_n = s_K(b-a)/h + 1.45$. The last result and equation (24) imply that

$$r_K T_n = r'_K T'_n + O_P(n^{-1/2} h^{-1}) \stackrel{a}{\sim} \chi_{a_n}^2.$$

C Conditions and Proof of Proposition 4

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

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

are a sequence of i.i.d. random variables, satisfying model (4) with i.i.d. 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 $\sigma^2(\cdot)$ 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 zero in such a way $nh \rightarrow \infty$.
- (B6) The function $F(x; \theta)$ is Lipschitz continuous in θ : $|F(x; \theta_1) - F(x; \theta_2)| \leq C(x)|\theta_1 - \theta_2|$, with $C(x)$ bounded in a neighborhood of m . In addition, the calibrated $\hat{\theta}$ is root- n consistent.

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

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

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

$$\hat{F}(m) - \bar{F}_0(m) = \hat{F}_c(m) - \bar{F}_c(m) + O_P(n^{-1/2}). \quad (25)$$

We now deal with the main term in (25). Write the local linear regression smoother as

$$\hat{F}_c(m) = \sum_{t=t_0-d}^{t_0+d} \sum_{i=1}^{N_t} W_{t,i}(m) \tilde{Y}_{t,i}, \quad (26)$$

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

$$\sum_{t=t_0-d}^{t_0+d} \sum_{i=1}^{N_t} W_{t,i}(m) = 1, \quad \sum_{t=t_0-d}^{t_0+d} \sum_{i=1}^{N_t} |W_{t,i}(m)| \leq 2 + o(1). \quad (27)$$

Then using (4) the data $\tilde{Y}_{t,i}$ can be written as

$$\tilde{Y}_{t,i} = \bar{F}_c(\bar{m}_{t,i}) + \varepsilon_{t,i} + Z_{t,i}, \quad (28)$$

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 ,

$$\sup_{\bar{m}_{t,i} \in \mathcal{N}} |Z_{t,i}| = O_P(n^{-1/2}). \quad (29)$$

Since K has a bounded support, all data points that contribute to computing (26) fall in \mathcal{N} . Therefore, in (26) replacing $\tilde{Y}_{t,i}$ by (28) and using (27) and (29), we have

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

where $\tilde{Z}_{t,i} = \bar{F}_c(\bar{m}_{t,i}) + \varepsilon_{t,i}$. It follows from (25) that

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

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

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

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

References

- Aït-Sahalia, Y., and J. Duarte, 2003, "Nonparametric option pricing under shape restrictions," *Journal of Econometrics*, 116, 9–47.
- Aït-Sahalia, Y., and A. W. Lo, 1998, "Nonparametric estimation of state-price densities implicit in financial assets prices," *Journal of Finance*, 53, 499–548.
- Amin, K., and V. Ng, 1993, "Option valuation with systematic stochastic volatility," *Journal of Finance*, 48, 881–910.
- Bakshi, G., C. Cao, and Z. Chen, 1997, "Empirical performance of alternative option pricing models," *Journal of Finance*, 52, 2003–2049.
- Bakshi, G., and Z. Chen, 1997, "An alternative valuation model for contingent claims," *Journal of Financial Economics*, 44, 123–165.
- Barone-Adesi, G., R. F. Engle, and L. Mancini, 2006, "A GARCH option pricing model in incomplete markets," NCCR FinRisk working paper, No. 155, University of Zurich.
- 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.
- Bekaert, G., and G. Wu, 2000, "Asymmetric volatility and risk in equity markets," *Review of Financial Studies*, 13, 1–42.
- Black, F., 1976, "Studies of stock price volatility changes," Proceedings of the 1976 Meetings of the American Statistical Association, Business and Economic Statistic Section, 177–181.
- Black, F., and M. Scholes, 1973, "The valuation 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., R. F. Engle, and D. B. Nelson, 1994, "ARCH models," in *Handbook of Econometrics*, ed. by R. F. Engle, and D. L. McFadden. North-Holland, Amsterdam, pp. 2959–3038.
- Campbell, J. Y., and L. Hentschel, 1992, "No news is good news: An asymmetric model of changing volatility in stock returns," *Journal of Financial Economics*, 31, 281–318.
- Carr, P., H. Geman, D. B. Madan, and M. Yor, 2003, "Stochastic volatility for Lévy processes," *Mathematical Finance*, 13, 345–382.

- Chernov, M., and E. Ghysels, 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.
- Christie, A., 1982, “The stochastic behaviour of common stock variances: Value, leverage and interest rate effects,” *Journal of Financial Economics*, 10, 407–432.
- Christoffersen, P., S. Heston, and K. Jacobs, 2006, “Option valuation with conditional skewness,” *Journal of Econometrics*, 131, 253–284.
- Comte, F., and E. Renault, 1998, “Long memory in continuous time stochastic volatility models,” *Mathematical Finance*, 8, 291–323.
- Constantinides, G., 1982, “Intertemporal asset pricing with heterogeneous consumers and without demand aggregation,” *Journal of Business*, 55, 253–268.
- Cont, R., and P. Tankov, 2004, “Non-parametric calibration of jump-diffusion option pricing models,” *Journal of Computational Finance*, 7, 1–49.
- Cox, J. C., and S. A. Ross, 1976, “The valuation of options for alternative stochastic processes,” *Journal of Financial Economics*, 3, 145–166.
- Derman, E., and I. Kani, 1994, “Riding on the smile,” *Risk*, 7, 32–39.
- Duan, J.-C., 1995, “The GARCH Option Pricing Model,” *Mathematical Finance*, 5, 13–32.
- Duffie, D., and C.-F. Huang, 1985, “Implementing Arrow-Debreu equilibria by continuous trading of few long-lived securities,” *Econometrica*, 53, 1337–1356.
- Dumas, B., J. Fleming, and R. E. Whaley, 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 I. Gijbels, 1995, “Data-driven bandwidth selection in local polynomial fitting: Variable bandwidth and spatial adaptation,” *Journal of the Royal Statistical Society, Series B*, 57, 371–394.
- Fan, J., and Q. Yao, 2003, *Nonlinear time series: Nonparametric and parametric methods*. Springer-Verlag, New York.

- Fan, J., C. M. Zhang, and J. Zhang, 2001, "Generalized likelihood ratio (GLR) statistics and Wilks phenomenon," *Annals of Statistics*, 29, 153–193.
- Fan, Y., and A. Ullah, 1999, "Asymptotic Normality of a Combined Regression Estimator," *Journal of Multivariate Analysis*, 71, 191–240.
- Fleming, J., B. Ost diek, and R. E. Whaley, 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., C. Gouri eroux, and E. Renault, 2005, "Efficient derivative pricing by extended method of moments," Working paper no. 2005-05, University of St. Gallen.
- Galai, D., 1983, "The components of the return from hedging options," *Journal of Business*, 56, 45–54.
- Ghysels, E., A. C. Harvey, and E. Renault, 1996, "Stochastic volatility," in *Handbook of Statistics*, ed. by G. S. Maddala, and C. R. Rao. North-Holland, Amsterdam, pp. 119–191.
- Glad, I. K., 1998, "Parametrically guided non-parametric regression," *Scandinavian Journal of Statistics*, 25, 649–668.
- Harrison, M., and D. Kreps, 1979, "Martingales and arbitrage in multiperiod securities markets," *Journal of Economic Theory*, 20, 381–408.
- Harvey, C. R., and R. E. Whaley, 1991, "S&P 100 index option volatility," *Journal of Finance*, 46, 1251–1261.
- 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 S. Nandi, 2000, "A closed-form GARCH option valuation model," *Review of Financial Studies*, 13, 585–625.
- Hull, J., and A. White, 1987, "The pricing of options on assets with stochastic volatilities," *Journal of Finance*, 42, 281–300.
- Madan, D. B., P. Carr, and E. Chang, 1998, "The Variance Gamma process and option pricing," *European Finance Review*, 2, 79–105.
- Melino, A., and S. Turnbull, 1995, "Misspecification and the pricing and hedging of long-term foreign currency options," *Journal of International Money and Finance*, 14, 373–393.
- Press, W. H., and J. W. Tukey, 1956, "Power spectral methods of analysis and their application to problems in airplane dynamics," Bell Telephone System Monograph 2606. Bell telephone System.
- Renault, E., and N. Touzi, 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., S. J. Sheather, and M. P. Wand, 1995, "An effective bandwidth selector for local least squares regression," *Journal of the American Statistical Association*, 90, 1257–1270.
- 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.
- Stutzer, M., 1996, "A simple nonparametric approach to derivative security valuation," *Journal of Finance*, 51, 1633–1652.
- Sundaresan, S., 2000, "Continuous-time methods in finance: A review and an assessment," *Journal of Finance*, 4, 1569–1622.
- Whaley, R. E., 1993, "Derivatives on market volatility: Hedging tools long overdue," *Journal of Derivatives*, 1, 71–84.
- 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.