# Fully Exploiting the Information Content of Intra Day Option Quotes: Applications in Option Pricing and Risk Management

Cameron Rookley

Department of Finance
University of Arizona

Previous Draft (April 1997):
(On the Use of Local Polynomial Regression to Smooth Intra Day Options Data and Estimate Risk Neutral Densities)

Revised
November, 1997

Abstract:

In this paper I present a data-intensive econometric smoothing method capable of fully exploiting the information content of intra-day option quotes. This is shown to be useful for both traditional option pricing applications as well as risk management. My method yields full sets of smoothed option prices with corresponding deltas and gammas for any "target time" of the day, while successfully addressing many common data problems, such as discontinuities in trading, the discreteness of option quotes and asynchronous trading. The procedures I outline here can be applied either in a real time setting or at the end of a trading day, so as to identify pricing errors or to estimate deltas and gammas for hedging purposes. In the process of smoothing the data, I also obtain reliable data-intensive estimates of the risk neutral density of the underlying asset at expiration. For the case of S&P 500 index options, these risk neutral density estimates are demonstrated to be superior to estimates suggested by the methods of Jackwerth and Rubinstein or Ait-Sahalia and Lo, and more importantly can reasonably be applied in practice.

## I. Introduction

Since the introduction of exchange-traded options in 1973, the financial world has seen a steady increase in the use of options and other derivative securities as a means of customizing state-contingent payoffs for both speculators and hedgers alike. Both the volume and the various types of option contracts available to investors has grown at a steady pace, with new types of exotic options being introduced from time to time. Given the wide-spread use of options in today's financial markets, the need to accurately price and hedge these contracts is self-evident. In this paper I present a data-intensive econometric model for estimating fitted option prices and common hedge parameters which fully exploits the information content of intra-day option quotes. My model uses local polynomial regression to smooth option prices with respect to their moneyness and the time of the day in which each quote occurs. By doing this, I successfully address many of the common problems often associated with high-frequency data such as discontinuities in trading, high degrees of noise and the potentially serious problem of asynchronous trading, where options with different strike prices and/or the underlying asset trade at different times during the day.

As well as serving as a useful pricing tool, my method is useful for risk-management as it provides reliable estimates of the volatility smile, common hedge parameters and the risk neutral density of the underlying asset for each unique option expiration date. Although of interest in their own right, the smile and risk neutral density are also used as inputs in other models. For example, starting with a set of risk neutral densities, Dupire (1994) provides an algorithm to recover a unique risk-neutral diffusion process consistent with observed (or fitted) option prices. Derman and Kani (1994) construct implied binomial trees from an observed volatility smile which is useful for pricing and hedging both standard and exotic options. Following the work of Stein and Stein (1991), Heston (1993) or more recently, Bates (1996), it is possible to link a non-standard risk neutral density to a rather flexible process driving the evolution of the underlying asset, allowing for possible jumps, stochastic volatility and correlations between the asset price and volatility. In all of these approaches the informational content of option prices is being exploited as a means of quantifying and possibly managing the risk associated with the random process of the underlying asset. Option based approaches to risk management are ideal as they

incorporate forward looking information into forecasts about future asset values. Options by there very nature reflect aggregate expectations regarding the likelihood of the underlying asset taking on various values at expiration. This approach to using the information content of option prices is in contrast to traditional models such as the CAPM, APT or time series methods such as GARCH, which generally attempt to quantify risk and/or expected returns using historical price data.

The use of option prices to infer certain properties of the underlying asset is not new. Lamoureux and Lastrapes (1993) demonstrate the value of at-the-money Black-Scholes implied volatilities in terms of helping to predict subsequent volatility. Also, the use of option prices to infer the risk neutral density of the underlying asset has a new and growing literature. Shimko (1993), Ait-Sahalia and Lo (1996), Jackwerth and Rubinstein (1996), as well as Masson and Perrakis (1997) consider estimating the risk neutral density for S&P 500 index options. Other asset classes include exchange rates: Campa, Chang and Reider (1997), Malz (1996), McCauley and Melick (1996a); interest rates: Abken (1995), McCauley and Melick (1996b); oil prices: Melick and Thomas (1997). Although several methods currently exist for estimating risk neutral densities, what makes this paper unique is the econometric method employed to accomplish the task. By using local polynomial regression with empirically driven bandwidths, I fully utilize the information content of a full day's worth of option quotes. This approach is shown to be very robust to noise and often yields reasonable results where other methods fails. As a result, the output from my model has practical applications for options pricing and risk management either in real time, at the end of a trading day, or for analyzing historical data.

The paper is organized as follows: in Section II I describe the use of implied volatilities for analysis, Section III introduces the semi-parametric model used to capture intraday option prices; in Section IV I discuss the nonparametric method used to estimate the model; Section V discusses the important issue of bandwidth choice; I present results for the case of S&P 500 index options in Section VI; simulation results are presented in Section VII and compared to recent work by Jackwerth and Rubinstein and Ait-Sahalia and Lo in Section IIX; Section IX concludes. For convenience, I relegate many of the details regarding data collection, detailed mathematical derivations and bandwidth choice into separate appendices.

## II Implied Volatilities

The seminal work of Black and Scholes (1973) provides a closed form solution for pricing a European call or put option on a non-dividend paying stock or asset. Modifying the formula for the case of predictable dividend streams is also straight forward. Although there is a preponderance of evidence to reject the assumptions underlying the Black-Scholes formula, the formula itself has evolved into a convenient mapping device for moving from an option's price and observed fundamentals to implied volatility. In fact, certain over-the-counter options on currency exchange rates are even quoted in terms of their implied volatilities. The fact that implied volatilities on options with different strike prices and expirations differ from each other, is evidence in itself that the Black-Scholes formula is incorrect, yet the use of implied volatilities for quoting the price of an option continues to be wide-spread. Regardless of the validity of the Black Scholes assumptions, the convention of quoting options in terms of implied volatilities is not without merit. The price of an option is in general a function of five observable variables and one unobservable. The variables which are typically treated as known include: strike price, price of the underlying asset, time to maturity, the risk free interest rate and expected dividend payments ($X,S_0,\tau,r,D$); the only unknown is the volatility of the underlying asset ($\sigma$). Given knowledge of the five knowns and the price of a European option, the implied volatility of the option is given by the value of $\sigma$ which equates the theoretical Black-Scholes price with the observed market price. The Black-Scholes formula for pricing a European call option with strike price $X_i$ , on a dividend paying stock or index with current level $S_t$, is given as follows:

$$C_{it} = (S_t - D)\Phi(d_1) - X_i e^{-r\tau}\Phi(d_2)$$

$$d_1 = \frac{\ln((S_t - D)/X_i) - (r + \frac{\sigma^2}{2})\tau}{\sigma\sqrt{\tau}} \qquad \textbf{(1)}$$

$$d_2 = d_1 - \sigma\sqrt{\tau}$$

Given a market price for $C_{it}$, the implied volatilities which are embedded in this formula prove to be indispensable tools for options pricing; they are a convenient measure of "price" as this is the

one item market participants are most likely to disagree upon. As the value of the underlying asset varies or the time to maturity begins to decay, the price of an option will correspondingly change; implied volatilities on the other hand tend to be less volatile and differences in implied volatilities convey much more economic information than option prices alone, as implied volatilities already embed much of the fundamental information available. Also, assuming the absence of riskless arbitrage opportunities, put-call-parity requires the implied volatility of a put and call option with the same strike and expiration be the same regardless of the validity of Black-Scholes. Finally, experience has shown that working with implied volatilities is much more convenient than attempting to explicitly model option prices and all of the various fundamental factors driving these prices directly. For all of these reasons, implied volatilities take center stage in the analysis which follows in the form of the volatility "smile".

The volatility smile - which plots implied volatility as a function of an option's strike price or moneyness - is the standard framework within which options are priced.[1] In a Black-Scholes world the implied volatility smile would be a flat line and the risk-neutral density log-normal; in practice many options such as S&P 500 index options have an implied volatility smile which tends to be a decreasing function of an option's strike price. As a result, in the money call options and out of the money puts tend to have prices which are above their theoretical Black-Scholes values, while out of the money calls and in the money puts are priced below their Black Scholes values. This volatility smile is empirically consistent with a risk neutral density which is not log-normal, and is the impetus for much of the current research on risk-neutral density estimation; in fact relatively high prices for out of the money puts is consistent with a risk neutral density with relatively long left tails or what Rubinstein refers to as possible "crash-o-phobia", as investors bid up the price of put options which serve as portfolio insurance.

**III. Smoothing in Implied Volatility Space**

As discussed above, many participants in the options market think and trade in terms of implied volatilities. Previous research also suggests working in implied volatility "space". For

[1]Moneyness is defined here in percentage terms and is given by the dividend adjusted asset price divided by the strike price for the case of a call option. So if a call option is "20% in the money", then its moneyness is 1.2.

example, starting with a set of closing prices, Shimko (1993) fits implied volatility as a quadratic function of strike price using a least squares criteria. Using a semi-parametric model, Ait-Sahalia and Lo (1995) use a year's worth of daily data to nonparametrically estimate implied volatility as a function of an option's strike price, the underlying index level and time to maturity. Jackwerth and Rubinstein (1996) start their analysis with every recorded quote for a particular trading day, and obtain a single set of option prices which are assumed to be representative of the day based upon the median implied volatility values for each unique strike and expiration. Even commercially available software products have features which deal directly with implied volatilities. For example, Option Pro by Essex Software Company allows the user to subjectively smooth an observed smile with the drag of a computer mouse. When working in this manner, the Black-Scholes formula simply serves as a mapping device for moving back and forth between implied volatility space and options price space.

In this paper, a method similar in nature to Ait-Sahalia and Lo is pursued with certain key differences. First, Ait-Sahalia and Lo use implied volatilities corresponding to a year's worth of daily option closing prices and smooth these values using Nadaraya-Watson type smoothers.[2] Although of academic interest for investigating the general characteristics of risk neutral densities, the frequency of this data is not very useful for practical pricing and risk management applications. In this paper intraday data is used one day at a time, so as to obtain estimates which are economically relevant for immediate practical use. Second, the choice of Nadaraya-Watson type smoothers can be improved upon by making use of local polynomial kernel smoothing; in fact the Nadaraya-Watson estimator used by Ait-Sahalia and Lo is a local polynomial kernel smooth of degree 0. By using higher order polynomial smoothing methods, superior estimates of the functions of interest can be obtained, especially at the boundaries. The boundary issue is particularly critical when estimating the tails of risk-neutral densities and for practical applications in real-time, as time of day is actually used as a smoother. Local polynomial kernel smoothing also provides a convenient and effective way to estimate the partial derivatives of a function of interest, which are necessary inputs for estimating option deltas, gammas, and risk neutral

[2]See Hardle (1989) for a discussion of traditional smoothing methods.

densities.

The particular approach I use is as follows: for a particular trading day and expiration, the implied volatilities corresponding to every available option quote are smoothed with respect to each option's moneyness (dividend adjusted index level divided by the strike price) and the time of day (minutes past 8:00 am) in which the quote occurred. This yields a synchronous set of fitted option quotes for any particular target time of interest. If applied in real time, the target time corresponds to either the time of the last observed quote or a relatively recent time interval. For each target time, I also estimate option deltas and gammas as well as the risk neutral density of the underlying index at expiration conditional on information available at that particular time of the day.

In order to apply local polynomial regression to a day's worth of implied volatility values, and achieve the goals outlined above, it is necessary to make certain assumptions regarding the evolution of the volatility smile across a single trading day. I assume implied volatilities for a given day and expiration are a smooth function of each option's current moneyness and the time of the day in which each option quote occurs. The following heteroscedastic model is therefore proposed to describe the nature of implied volatilities:

$$\begin{aligned} \sigma_{it}\,(\,M_{it}\,,\,t\,) &= \mu\,(\,M_{it}\,,\,t\,) + \sqrt{h_{it}}\;\epsilon_i \\ E(\epsilon_i) &= 0 \quad , \quad Var(\epsilon_i) = 1 \\ h_{it} &= g\,(\,M_{it}\,,\,t\,) \end{aligned} \qquad \textbf{(2)}$$

where implied volatility ($\sigma_{it}$) is the dependent variable, $\mu(\,)$ represents the conditional mean which I assume is a smooth function of moneyness and time ($M_{it}$ and t), and $\epsilon_{it}$ represent I.I.D error terms. $h_{it}$ represents the conditional variance of $\sigma_{it}$ which is I also assume is a smooth function of $M_{it}$ and t . By smoothing with respect to moneyness only ($M_{it} = (S_t\text{-}D)/X_i$ ), as opposed to the dividend adjusted index level and strike price as two separate variables, I implicitly assume the theoretical option function is homogenous of degree one with respect to the index and strike price. The basic Black-Scholes formula is an example of such a function, and as shown by Merton (1973) and discussed in Ingersoll (1987), a call price is homogenous of degree one in the

asset price and strike price if the asset's return distribution is independent of the level of the underlying index. For example, the pricing formula associated with a stochastic volatility model where the drift and diffusion functions depend on the volatility itself but not on the asset price would satisfy this assumption. Ait-Sahalia and Lo, in order to avoid possibly restrictive assumptions on the diffusion, choose to model implied volatility as a function of the index, strike price, and time to maturity separately. In my setup, time to maturity is essentially held constant since we deal with a single day at a time. Furthermore, although a single day's worth of quotes can potentially consist of several thousand observations for each expiration date, I choose to use moneyness as a single smoother so as to limit the curse of dimensionality associated with adding an additional variable[3]. Moreover, my use of time of day as an additional smoother can potentially capture any effects associated with changes in the returns distribution due to changes in the asset price or other factors over time, and hence compensate for any errors introduced by using moneyness as a single smoother. Finally, the impact of any incremental decreases in the time to maturity of the option over the course of a single trading day is likely to be picked up by this variable.

Since we are assuming option prices are homogenous to degree 1 with respect to moneyness, it is convenient to introduce the following notation which standardizes option prices by dividing them by their dividend adjusted index levels:

$$c_{ti} = \left(\frac{C_{ti}}{S_t - D}\right) \quad , \quad M_{ti} = \frac{(S_t - D)}{X_i} \quad , \quad c_{ti} = f(M_{it}, \sigma(M_{it}, t)) \qquad \textbf{(3)}$$

where $C_{ti}$ ($c_{ti}$) represents a (standardized) call price at time t with strike price $X_i$ ,$S_t$ represents the underlying index at time t, D represents the present value of all dividends paid from now until maturity, and $M_{it}$ represents the moneyness of an option with strike $X_i$ at time t. For the

[3]The curse of dimensionality is related to the sparseness of data which is often encountered with higher dimensions; as a result rates of convergence tend to decrease significantly with the introduction of one or more smoothers when conducting local polynomial regression.

semiparametric model which follows, the function f(.) represents the Black-Scholes formula divided by the dividend adjusted index level ($S_t$-D), and σ(.) represents implied volatility which is modeled as a function of each option's moneyness and the time of day of the quote. Specifically:

$$c_{it} = f(M_{it},\sigma(M_{it},t)) = \Phi(d_1) - \frac{e^{(-r\tau)}\Phi(d_2)}{M_{it}}$$

$$d_1 = \frac{\ln(M_{it}) + (r + \sigma(M_{it},t)^2/2)\tau}{\sigma(M_{it},t)\sqrt{\tau}} \qquad \textbf{(4)}$$

$$d_2 = d_1 - \sigma(M_{it},t)\sqrt{\tau}$$

where Φ represents the normal C.D.F. Although the standard Black-Scholes model requires σ be constant across M and t, the model given above can accommodate non-horizontal volatility smiles as well as changes in the overall level or shape of the smile which tend to occur even during the course of a single trading day. The framework outlined above is designed to capture this phenomena without imposing any assumptions beyond certain smoothness conditions on the nature of the relationship between σ, M and t. Furthermore, letting $V_{it}=\sigma(M_{it},t)$, the following two derivatives of interest can be estimated:

$$\frac{\partial V_{it}}{\partial M_{it}} \quad , \quad \frac{\partial^2 V_{it}}{\partial M_{it}^2}$$

These derivatives are crucial for estimating hedge parameters risk neutral densities. I use the first partial derivative of the volatility surface with respect to moneyness in conjunction with the Black Scholes formula to estimate option deltas (Δ=dC/dS), and the second partial derivative of volatility with respect to moneyness to estimate option gammas (γ=dΔ/dS) and the risk neutral density of the underlying asset at expiration. The use of this second derivative to arrive at the risk neutral density of the index relies upon the Breeden and Litzenburger (1978) result. If options are priced in the risk-neutral measure, then the risk neutral density of the index at expiration

$\tilde{f}(S_T)$ is implicitly embedded in the risk neutral valuation formula for a European call. By taking derivatives of this formula with respect to strike price we obtain:

$$\begin{aligned} C &= e^{-r\tau}\int_X^{+\infty}(S-X)\tilde{f}(S)dS \\ \frac{\partial C}{\partial X} &= e^{-r\tau}\int_X^{\infty}-\tilde{f}(S)dS = e^{-r\tau}(F(X)-1) \\ \frac{\partial^2 C}{\partial X^2} &= e^{-r\tau}\tilde{f}(X) \end{aligned} \qquad \textbf{(6)}$$

This follows from a direct application of Liebnitz rule. Therefore by calculating the second derivative of the call price with respect to the strike, we gain knowledge of the risk neutral density of the index at expiration evaluated at X[4] This second derivative calculation, as well as the formulas for $\Delta$ and $\gamma$ are tedious yet straight forward to calculate in the context of the semi-parametric model outlined above; Appendix A contains the details of these calculations.

**IV. Local Polynomial Kernel Smoothing**

In order to estimate $\sigma(M,t)$ and its associated first and second derivatives with respect to M, I use bivariate local polynomial kernel regression[5]. The basic idea of local polynomial regression rests upon the use of locally weighted least squares regression, where individual weights are determined by three factors: the choice of a kernel function, the distance of an observation from a certain "target point" and the chosen bandwidth vector. In practice the choice of the kernel function is of minor importance relative to the choice of the bandwidth vector. For computational efficiency, a multiplicative Gaussian kernel is used with an empirically chosen local bandwidth. The details of this bandwidth choice procedure are delayed until the next section as well as Appendix C.

---

[4]Note also that we can obtain the CDF of the risk neutral density by observing the first derivative of the call price with respect to the strike.

[5]Wand and Jones, *Kernel Smoothing*, 1995, is an excellent source for this topic.

In practice kernel smoothing can be extremely computationally expensive. Fortunately for this application, we only need to estimate the two-dimensional implied volatility surface and it's associated derivatives at a limited set of target points. This is in contrast to many applications where fitted values are sought for every single observation point in the raw data. To obtain a full set of synchronous smoothed and fitted options prices corresponding to a particular target time of the day, I select a grid of moneyness points which correspond to the underlying dividend adjusted index level at that target time, divided by each unique strike price. For density estimation we can choose a single target time such mid-day or end-of-day and evaluate the volatility surface and its derivatives over a fine grid of moneyness points corresponding to this time. Using the Breeden-Litzenburger result and the formulas contained in Appendix A, it is then possible to estimate the risk-neutral pdf of interest over this fine grid of points.[6]

To understand the mechanics of local polynomial kernel smoothing, let $(M_{i\tau},\tau)$ represent a target point of interest. In order to estimate the function at this point, we fit a local polynomial using nearby observations which are weighted by their distance to $(M_{i\tau},\tau)$. Given a trivariate data set consisting of $\{M_j,t_j,\sigma_j\}$, j=1,....n, . a fitted value corresponding to this target point is given as follows:

[6]This is a further advantage of smoothing with respect to moneyness as opposed to strike prices directly. The smoothing procedure allows us to evaluate the function at fictional strike prices for a given index level. When calculating derivatives we simply make use of the fact that $X_i=(S_t\text{-Div})/M_i$, where $S_t$ is the index level at the target point.

$$\hat{\sigma}(M_{i\tau},\tau) = e'\hat{\beta}_i = e'(X_i' W_i X_i)^{-1} X_i' W_i Y$$

$$e'=(1\ 0\ .....\ 0)$$

$$X_i = \begin{matrix} 1 & (M_1-M_{i\tau}) & \dots & (M_1-M_{i\tau})^{P_M} & (t_1-\tau) & \dots & (t_1-\tau)^{P_t} \\ 1 & (M_2-M_{i\tau}) & \dots & (M_2-M_{i\tau})^{P_M} & (t_2-\tau) & \dots & (t_2-\tau)^{P_t} \\ .... & .... & & .... & ... & & \\ 1 & (M_n-M_{i\tau}) & \dots & (M_n-M_{i\tau})^{P_M} & (t_n-\tau) & \dots & (t_n-\tau)^{P_t} \end{matrix} \quad \textbf{(7)}$$

$$diag(W_i) = (\phi((M_1-M_{i\tau})/h_M)\cdot\phi((t_1-\tau)/h_t),\dots,\phi((M_n-M_{i\tau})/h_M)\cdot\phi((t_n-\tau)/h_n))$$

$$Y'=(\sigma_1,\sigma_2,..,\sigma_n)$$

The estimate of σ at the target point is therefore the constant term from a WLS regression where the weights along the diagonal of W are given by a multiplicative Gaussian kernel $\phi((M-M_\tau)/h_M.)\phi((t-\tau)/h_t)$. Larger values of either bandwidths ($h_m$ and $h_t$) lead to higher weights being attached to observations which are farther away as ϕ represents the Gaussian function. Therefore by increasing the values of $h_m$ or $h_t$ it is possible to obtain increased levels of smoothing and indeed the choice of the bandwidth vector $h_m|h_t$ is the key decision to make. The constants $P_M$ and $P_t$ represent the degree of the polynomial in each direction. Setting $P_M=P_t=1$ produces the special case of local linear fitting; this amounts to locally fitting a hyerplane through each target point using kernel weighted observations in a two-dimensional neighborhood around that point.[7]

The framework described above is easily adaptable to derivative estimation as well. For example in order to estimate the r'th derivative of σ with respect to moneyness, the $(1+P_M+P_t)$x1 vector e=(1,0....0) used above is replaced with a $(1+P_M+P_t)$x1 vector $e_r$ which is a vector of zeros

[7]When using higher order polynomials, cross product terms are often introduced as well. For example, if $P_m=P_t=2$ it would not be uncommon to include a term such as $(M_j-M_{i\tau})(t_j-\tau)$ as an added regressor.

except for a single 1 corresponding to the r'th polynomial term of $(M_j - M_{i\tau})$. The r'th derivative is then given by r! multiplied by the same formula used above for estimating the function in levels, with the only other modification being the vector e'. If for example $P_m = P_t = 2$, estimates of the first and second derivative with respect to Mit are given by:

$$\frac{d\hat{\sigma}_{M_{it},\tau}}{dM_{i\tau}} = e_1'(X_i'W_iX_i)^{-1}X_i'W_iY$$

$$\frac{d^2\hat{\sigma}_{M_{it},\tau}}{dM_{it}^2} = 2e_2'(X_i'W_iX_i)^{-1}X_i'W_iY \qquad \textbf{(8)}$$

$$e_1' = (0,1,0,0,0) \quad , \quad e_2' = (0,0,1,0,0)$$

where $X_i, W_i$ and Y are defined as before. Therefore the first derivative estimate at the target point is the slope coefficient in the M direction, and the second derivative estimate is two times the second order coefficient in the M direction. This follows from differentiating the function for fitted implied volatility values either once or twice with respect to $M_j$ and evaluating the function of interest at $M_j = M_{i\tau}$. Note from (8) that unless one wishes to use simple difference methods based upon a lower order derivative, it is necessary that the degree of the polynomial be greater than or equal to the order of the derivative being estimated.

**VI. Bandwidth Choice**

Probably the most critical decision to make in the context of local polynomial fitting is the choice of the bandwidth vector. This decision involves a classic bias-variance trade off, where the larger the bandwidth the more the function is smoothed. If a bandwidth is chosen to be too low, the estimated function tends have low bias but high variance. If the function is over-smoothed, variance can be reduced, but only at the cost of introducing higher bias. The three panels of Figure 1 (attached) represent a local linear fit of the implied volatility surface for S&P 500 Index Options on January 4, 1993, with a March expiration date. The raw data originates from the Berkeley Options Database as well as other sources. The details of this database, as well as the steps I take to arrive at implied volatilities and other necessary variables have been relegated to Appendix B. Candidate bandwidth vectors are given by $h_M | h_t$ in the titles above each panel.

Figure 1-A involves the least amount of smoothing and as a result the surface is somewhat erratic. Moving from Figure 1-A to Figures 1-B and 1-C, I introduce more and more smoothing. These diagrams clearly illustrate the importance of bandwidth choice. Under-smoothing leads to a surface such as the one in Panel A, while over smoothing runs the risk of removing possibly important features of the data. The possible consequences of over smoothing are best illustrated in Figure 2. Figure 2 graphs the corresponding mid-day risk neutral pdfs, as derived from the surfaces represented in Figure 1. In order to arrive at these pdfs, I estimate the first derivative of the surface with respect to moneyness using the same global bandwidth as for the surface. I obtain the second derivative by first differencing the first derivative estimate. Combining these quantities with the formulas in Appendix A we arrive at the risk neutral density. In Figure 2, I also plot a base-line log-normal distribution corresponding to a volatility value of $\sigma$=11.5%, which is the median implied volatility across all quotes for this day and expiration.[8] Several features are worth noting from Figure 2. The actual range of available strikes is 375 to 460; Figure 2 therefore illustrates the ability of local polynomial smoothing to not only perform well at the boundaries, but also to extrapolate into areas which are absent of data as evidenced by the smooth tails of the estimated distribution when h=0.05|25 or higher. The second feature worth noting is the somewhat erratic behavior in the tails of the distribution when using the relatively low bandwidth vector $h_M|h_t$=0.025|12.5. These anomalies are removed as I increase the size of the bandwidth, but by doing so, we run the risk of over smoothing the data; in this case a high bandwidth appears to force the estimated risk neutral distribution towards the standard Black-Scholes log-normal distribution. Finally, note the difference in shape between the base-line log normal distribution corresponding to the standard Black-Scholes model and the empirically derived risk-neutral distributions. This non-trivial difference highlights the importance of seeking alternatives for risk-neutral density estimation as well as the implications of a non-constant smile on the risk-neutral pdf; in particular relatively high implied volatility values for low strikes, translates into a risk neutral pdf which is skewed to the left.

[8]Whether we used median implied volatilities across all quotes, mean implied volatilities across all quotes or mean or median implied volatilities for at the money options, made little difference, as all 4 of these values were within half a percentage point of 11.5%.

It should be stressed that I estimated the implied volatility surfaces of Figure 1 using the same bandwidth vector for each target point in the surface. In practice we are not restricted to using a single global bandwidth and several factors should influence the choice of the bandwidth at a particular point. These factors include: the number of observations in the neighborhood of the target point, the variance and curvature of the true underlying function at that point, and the order of the function being estimated.

So the question remains: How much should the data be smoothed? There are two basic approaches to addressing the smoothing question. The first approach relies on interactive graphing techniques where we start with a low bandwidth and gradually increases the bandwidth until the function is sufficiently smooth to the appearance of the eye. Figure 1 is an example of this approach. The approach itself is inherently subjective but often valid depending upon the end purpose of the research in question. The other general approach is to rely upon automatic bandwidth selectors. These methods attempt to minimize a global criteria such as asymptotic mean integrated squared error or a cross-validation function.[9]

There is a tremendous amount of literature in the area of automatic bandwidth selection techniques, with most of the discussion to date focusing on the univariate case; one notable exception to this is a recent paper by Ruppert (1997) which uses a method known as empirical bias bandwidth selectors (EBBS). This method of bandwidth selection, although computationally expensive, is unique in its ability to estimate data-driven **local** bandwidths for multivariate functions as well as any order derivative of interest. In the process of implementing the EBBS approach it is also possible to obtain estimates of the heteroscedastic error function for both the underlying function and function derivatives. The general approach of this method is to empirically estimate both the bias and variance of the function of interest, at each target point, using several candidate bandwidths. Bandwidths are then chosen locally so as to minimize the

[9]Cross validation attempts to minimize the sum of squared errors between smoothed and actual values of the dependent variable. Notice that an obvious unrestricted choice would be to simply take actual values of the dependent variable as the smoothed values. To rule out this under smoothing scenario, cross validation estimates fitted values at each point without actually using the data point corresponding to the actual value. For this reason cross-validation functions are often referred to as “leave-one-out estimators”.

estimated mean squared error at each point. The exact details of this method can be found in Ruppert (1997), as well as Ruppert, Wand and Hossjer (1997). In Appendix C I provide a general outline of the procedure, as well as the various "tuning" parameters which must be chosen to implement the method in practice.

In Figure 3 I provide an example of this method using S&P 500 index option quotes for January 4, 1993 and a March expiration date. Figure 3-A graphs the estimated implied volatility surface using the EBBS bandwidth approach, while Figure 4-A illustrates the square root of the heteroscedastic variance function. To the right of these figures (Figures 3-B and Figures 4-B), are the optimal local bandwidths I use to estimate each function. When using the EBBS approach the data is first standardized so that I use a common local bandwidth in both directions; the local bandwidth presented in these graphs is therefore in terms of the standard deviation of the data in each direction. In Figure 5, I present similar results for the estimation of the first derivative with respect to the variable moneyness.

**VII Results**

In this section I illustrate the shape of some estimated implied volatility surfaces with their corresponding risk-neutral distributions for January 4, 1993. The 4 panels of Figure 7-A represent implied volatility surfaces for the nearest 4 expiration months. I estimate these surfaces using local quadratic fitting with optimal bandwidths chosen by EBBS. The first and second derivatives with respect to moneyness are also estimated using local quadratic fitting, but the same optimal bandwidth for estimating the surface is used for derivative estimation as well. In Section VIII, which follows, I discuss the advantages of using a common bandwidth for estimating all 3 of these functions.

Referring to the four panels of Figure 7, note the commonly observed volatility smile where implied volatility values tend to be higher for in the money call options as well as deep out of the money calls; also note how the shape of the volatility smile does tend to evolve, even during the course of a single trading day; this appears to be especially true for the nearer expiration dates. The four panels of Figure 8 plot the estimated risk neutral distributions corresponding to these surfaces. In each graph the solid line represents the pdf as derived from the smooth as well as a parametric match for each distribution. As shown by Johnson and Kotz (1970) and used by

Longstaff (1995) for risk neutral density estimation, the following 4-parameter edgeworth expansion type density can match the first four moments of any continuous density:

$$z = \frac{\ln(S_T) - \alpha}{\sigma}$$

$$f(S_T) = \left(\frac{1}{S_T \sigma}\right)\phi(Z)(1 + \beta(z^3 - 3z) + \gamma(z^4 - 6z^2 + 3)) \quad \textbf{(10)}$$

This density allows for a wide variety of possible shapes, of which the log-normal is a special case ($\beta$=0, $\gamma$=0). Since the estimates of the risk neutral pdf as derived from the smooth are based upon a limited range of strike prices, it may be advantageous to parameterize such densities as a means of extrapolating into the tails or for estimating moments. Although it is possible to extrapolate beyond the actual range of the data using nonparametric regression, smoothing methods tend to perform quite poorly if one attempts to extrapolate too far. By parameterizing the distribution, it is possible to force the tails of the estimated distribution to be relatively well-behaved while guaranteeing the parameterized distribution integrates to one. I estimate the 4-parameters ($\alpha$,$\sigma$,$\beta$,$\gamma$) for these parametric density "matches" by minimizing the mean integrated squared error between the parametric pdf and the pdf from the smooth, over the actual range of the data. In Figure 9 I combine these 4 parametric distributions into one diagram so as to illustrate the increased uncertainty associated with the stock index, as expiration dates further away are considered.

In this section I illustrate the use of the semiparametric model in practice which appears to yield reasonable results. Unfortunately without true knowledge of the data generating process it is difficult to truly assess how well the model is doing; the simulations of the next section address this issue.

**VII Simulations**

In order to evaluate the accuracy of my method and make comparisons to related work, it is necessary to conduct simulations where the actual model is known with certainty. To do this several I simulate several "true" implied volatility surfaces with noise. I then estimate the various functions of interest using the methods described above, thereby allowing the performance of the

bandwidth selection procedure and the use of different ordered polynomials to be assessed. The robustness of my method to different levels of noise is also analyzed with a focus on the accuracy of estimated implied volatilities, deltas and gammas, as well as the accuracy of risk-neutral density estimates. I pay special attention to estimation near the time boundary. Accurate estimation near the time boundary is particularly important if fitted values are to be used in real time, for the purpose of identifying pricing errors or for estimating the risk neutral density conditional upon recently available information.

In order to begin the simulations it is necessary to decide upon a true model. It is also desirable to have a "true" model which might reasonably be encountered in practice. To accomplish this goal, I parametrically fit implied volatility surfaces from the actual data and use these as candidate "true" surfaces to simulate from. I estimate these surfaces using generalized least squares, where actual implied volatilities are regressed on the 9 variables which comprise the second-degree polynomial of moneyness and time, including interaction terms. I also allow for possible heteroscedasticity with respect to the moneyness of each option as well as time of day. Specifically:

$$\sigma_{it} = X_{it}\beta + e_{it}$$

$$E(e_{it}^2) = Z_{it}\gamma$$

$$X_{it} = (1 \; M_{it} \; M_{it}^2) \otimes (1 \; t \; t^2)$$

$$Z_{it} = (1 \; M_{it} \; M_{it}^2 \; t)$$

Since the use of this second order polynomial can often lead to very erratic fitted surfaces, it is desirable to rule out such surfaces as a "true model" for the purpose of the simulations which follow; I therefore discard any estimated surface which has a corresponding risk neutral density which is not unimodal at every point in time for the day or contains negative values. I apply the above procedure to every unique trading day and expiration date for January 1993, which yields a set of 30 volatility surfaces to simulate from out of a possible universe of 92. The fact that I discarded more than two thirds of the candidate risk neutral pdfs as implausible highlights the

poor performance of parametric polynomial fits (such as the method suggested in Shimko (1993)) in terms of generating reasonable risk neutral pdfs. In contrast local polynomial regression as employed in this paper, tends to always generate reasonable risk neutral densities in practice. For each candidate surface, I use actual values of moneyness and time of day to generate true implied volatility values as $Z_{it}b$, and attach a random normal error term with mean 0 and variance $Z_{it}g$ to each observation, where b and g are estimates of $\beta$ and $\gamma$ in the GLS model given above. After generating the data, I estimate the implied volatility surface and its associated derivatives over a Cartesian grid with 20 equally spaced moneyness points, and 7 hourly target times from 9:00 am to 3:00 pm.[10] I then use estimates corresponding to the 3:00 pm target time for evaluating the performance of estimation near the time boundary.

Given estimates of the implied volatility values with corresponding deltas, gammas and risk neutral pdfs at the target points, I compare these estimates to their true values using a mean absolute deviation (MAD) and root mean squared error (RMSE) criteria. For calculating actual and fitted risk neutral pdfs, I assume the actual index level is equal to the median index level for the day. For each of the 30 "true" surfaces, I run ten simulations by drawing a new vector of realized error terms with each run. Instead of conducting the EBBS approach for every run, I use optimal local bandwidths corresponding to the first simulation run for each candidate surface; this is justified as the fundamental factors which affect local bandwidth choice are the same with each new draw. The results which follow are therefore based on a total of 300 simulations.

Using the box plots in Figures 10 and 11, I illustrate the performance of different ordered polynomial schemes in terms of estimating the various functions of interest These polynomial schemes differ in terms of the polynomial which is locally fitted and whether or not a common optimal bandwidth is used for estimating the surface as well as the first and second derivatives with respect to moneyness. For the case of P=0, I fit the surface using Nadaraya-Watson type smoothing with a bandwidth chosen by EBBS; I then obtain $d(i.v)/dm$ and $d^2(iv)/dm^2$ with simple differencing. For $P=1_O$, local linear fitting is used to estimate the surface as well as d(i.v.)/dm,

[10]The first option quote for each day, usually occurs within a few minutes of 8:30 am, and the final option quote is usually within a few minutes of 3:15 pm. The target grid of moneyness points begins at the actual minimum value and extends to the actual maximum value.

with optimal bandwidths for both chosen by EBBS; $d^2$(i.v.)/dm2 is then obtained by first differencing d(i.v.)/dm. $P=1_C$ is the same as $P=1_O$ except I use a common optimal bandwidth to fit the surface as well as the first derivative. Finally, $P=2_O$ and $P=2_C$ refer to local quadratic fitting using separate optimal bandwidths for all three functions of interest or a common bandwidth for all three.

Based on these box-plots, local quadratic fitting appears to be overall superior in terms of estimating the functions of interest relative to lower ordered polynomials. Furthermore, whether I select optimal bandwidths for all three underlying functions or whether I use a common bandwidth appears to have little effect on the final results. This last observation turns out to be very advantageous for two reasons: first, the EBBS approach is computationally expensive and searching for a single optimal local bandwidth for all three functions saves considerable amounts of computing time; secondly, using a common bandwidth for all three underlying functions makes it possible to estimate a variance covariance matrix for the three functions at each target point in the estimation.[11] With an estimate of the variance covariance matrix it is straightforward to simulate realizations of more complicated functions such as option deltas, gammas and the risk neutral pdfs which depend upon $\sigma(M_{it},t)$, $(d\sigma/dM_{it})$, and $(d^2\sigma/dM_{it}^2)$. The output of these simulations can then be used to construct confidence bounds for these highly nonlinear functions. The four panels of Figure 12 illustrate local quadratic fits of the volatility smile, option deltas and gammas, as well as the risk neutral distribution for mid-day January 4, 1993, using March expiration date options. I Include with each graph 95% confidence bounds based upon fitted values plus and minus 2 estimated standard deviations. For the case of option deltas, gammas and the risk neutral distribution, I make 10,000 draws using fitted values for each function and the estimated variance covariance matrix for $\sigma(M,t)$, $d\sigma(M,t)/dm$, and $d^2\sigma(M,t)/dM^2$ at each target point. Not surprisingly, the accuracy of the estimated functions tends to diminish as we get further in or out of the money towards the boundary points in terms of the moneyness variable. With reference to Figures 10 and 11, whether I use local quadratic fitting or not, estimation performance at the time boundary is significantly worse as well. This is a common feature of

[11]See Equation 34 in Appendix C for details on this variance covariance matrix.

nonparametric regression and raises an interesting question for practically implementing this method on a real-time basis. Should the risk neutral distribution and other functions be estimated at day's end using a target time corresponding to the last available quote, or should the target time be set before the end of day perhaps one or two time bandwidths away from the last quote? In practice it might be advisable to estimate both and see if these functions significantly deviate from each other. Finally, if the purpose of the research is to obtain a single risk neutral density estimate so as to represent the day as a whole, (as in Jackwerth and Rubinstein (1996)) it would clearly be advisable to choose an interior target time for this purpose. In the section which follows we compare the method proposed in this paper to other methods of risk neutral density estimation.

**IX Comparison With Other Methods**

As discussed in the introduction, several methods currently exist for estimating risk neutral densities. For the case of S&P 500 index options, Shimko(1993), Ait-Sahalia and Lo (1997), Rubinstein (1994), Jackwerth and Rubinstein (1997) and Longstaff (1995) all propose different methods of estimating risk neutral densities. All of these alternative methods suffer from certain drawbacks however. Shimko's use of simple quadratic fits is often inadequate in terms of capturing salient features of the data; at the same time the use of higher order polynomials for fitting a single set of implied volatilities will frequently lead to implausible pdfs in practice. Ait-Sahalia and Lo's semi-parametric model is similar in nature to the model presented here, but the purpose of the former is to estimate an option pricing *function* which is assumed to hold across different calender dates and expirations. Although effective in terms of identifying general features of the data, as a pricing tool, the method fails to correctly price all options at all points in time, and as noted by the authors "there may be some dates for which the SPD fits the cross section of option prices poorly and other dates for which the SPD performs very well."[12] By contrast the model I present here is designed to be relevant for immediate practical use, and allows the data generating process to change with time so that the volatility smile and corresponding risk neutral density reflects any concurrent information pertaining to aggregate

[12]Ait- Sahalia and Lo, (1997), p. 13.

expectations regarding the future. From an econometric standpoint, the choice of Nadayara-Watson type smoothers using optimal *global* bandwidths can probably be improved upon as well - particularly for estimation near the data boundaries - with the use of higher ordered local polynomials and optimal *local* bandwidths.

Unlike Ait-Sahalia and Lo, the methods of Rubinstein (1994) and Jackwerth and Rubinstein (1997) are designed to be used with a single cross section of current option prices. Rubinstein's (1994) method relies upon the use of a "prior" type density, and estimates points of discrete mass under the risk neutral distribution (state prices) so as to minimize a chosen distance measure between the estimated and prior distributions while forcing all assets to be priced correctly within their bid-ask spreads. Jackwerth and Rubinstein (1997) discuss the use of various prior's and distance measures, and also suggest an alternative method for estimating the risk neutral distribution which relies upon minimizing a penalized smoothness criteria. The idea behind penalized smoothness is to nonparametrically estimate a smooth density which penalizes departures from smoothness, as well as departures of estimated option prices from their actual values.

The biggest drawback of these last two methods is their inability to accommodate non-trivial levels of noise. The pricing constraints in Rubinstein's (1994) method require fitted option prices to fall within actually observed bid-ask spreads, however in a recent study by Masson and Perrakis (1997) the authors were unable to find a single candidate distribution capable of satisfying these constraints in more than 2/3 of the observed sets of option quotes they looked at.[13] Although Jackwerth and Rubinstein's (1997) penalized smoothness criteria is not as restrictive with respect to the pricing constraints, this recent attempt proves to be non-robust to noise as well; in fact the procedure of aggregating a day's worth of option quotes using median implied volatility values for each unique strike price often generates more noise than their method can reasonably handle.

Using the simulation environment of Section VII it is possible to illustrate the performance of the penalized smoothness method in the face of noise levels which might reasonably be encountered in practice. Figure 13 illustrates the potential pricing errors which can occur by

[13]Masson, J. And Perrakis, S., "A Jumping Smile", working paper, March 1997.

aggregating option quotes using median implied volatility values. Included in this graph is the actual simulated volatility smile corresponding to mid-day, with 95% confidence bounds. Also plotted is a volatility smile corresponding to median implied volatility values for each strike, and the implied volatility values corresponding to each strike price which occur as close as possible to the mid-day target time. This graph corresponds to the first simulation and comparable results are obtained when simulating from the other 29 surfaces as well. What Figure 13 illustrates, is the inability of median implied volatility values to successfully aggregate away the noise which is encountered in practice. In this case, whether we use closest time values or median values to obtain a synchronous set of quotes, the variation of the smile from it's actual values is substantial and crosses the 95% confidence bounds for both in this particular case.

Given this level of noise, how does the penalized smoothness criteria perform? Using the median implied volatilities of Figure 13, I convert these values to option prices using Black-Scholes and use these prices as inputs for estimating the risk neutral pdf. Specifically I use Jackwerth and Rubinstein's penalized smoothness criteria as an objective function with several possible penalty parameters.[14] Using the terminology of Jackwerth and Rubinstein, the set of possible index prices is then "clamped down upon" until all values for the pdf are non-negative. Figure 14 plots the "true" and estimated pdfs for the first simulation using this method. For the relatively larger values of $\alpha$, the estimated pdfs are bimodal, sharply truncated to the right, and have most of their mass centered to the right of the actual mean. As we reduce $\alpha$ and discount the information content of the option quotes further, we manage to remove the bimodality of the estimated distribution but continue to have sharply truncated distributions which also appear to overestimate the mean. This particular example illustrates the inability of Jackwerth and Rubinstein's method to reasonably recover an actual risk neutral pdf, given a level of noise likely to be encountered in practice. The assumption of a constant volatility smile for each trading day is one potential reason for this poor performance. In fact, Jackwerth and Rubinstein acknowledge this shortcoming and suggest filtering the data so as to remove trading days in which the volatility smile exhibits significant intraday movements. For the trading days and

[14]In this case lower levels of $\alpha$ allow greater departures of actual option values from their fitted values, and this in turn leads to smoother risk neutral pdfs.

expiration dates considered in their paper, the filter employed leaves 1172 out of 1850 sets of option quotes to conduct analysis on.[15]

Given that the smile and associated pdfs of the previous example (Figures 13 and 14) may be associated with unusually high levels of noise due to a large daily movement in the index, asynchronous trading, or other possible data problems, I also investigate the performance of the penalized smoothness method in the presence of moderate well-behaved noise terms. Specifically, I add a random normal error term with standard deviation of 0.01 to actual midday implied volatility values; and once again convert these implied volatility values into option quotes for pdf estimation. In Figure 15 I illustrate the performance of penalized smoothness for this example; for relatively large values of $\alpha$ the pdf estimates appear to be relatively close to the actual density; yet they are either too curvy near the shanks of the distribution or over smoothed for the case of low values of alpha. Once again the penalized smoothness method appears to perform poorly, even in the presence of well-behaved error terms. Using Figures 13-15, I demonstrate the sensitivity of Jackwerth and Rubinstein's method to noise, and this feature of their method presents a major obstacle for implementing it in practice. In contrast to Jackwerth and Rubinstein's method, the semiparametric local quadratic fits presented in this paper are very robust to noise and can successfully estimate a plausible risk neutral pdf virtually every time in simulations, as well as in practice. If we further parameterize these distributions by matching them with an edgeworth expansion type density, it is also possible to obtain a density which is well-behaved in the tails and we can obtain sample moments which depend upon the entire density.

This last point raises an interesting question though. Why don't we attempt to estimate the edgeworth expansion type density directly as in Longstaff (1995)? For example, it is possible to estimate the four parameters of this density so as to price four options exactly, or to minimize some objective function which depends upon a full cross section of option prices. Somewhat surprisingly though, direct attempts at parametrically estimating risk neutral densities such

---

[15]My method, on the other hand, models intraday movements in a direct fashion and the use of such a filter is clearly unnecessary making the method presented here more widely applicable in practice.

edgeworth expansion type densities, or mixtures of log-normals, usually lead to implausible distributions which are multi-modal and/or contain negative values over the support of the pdf. This is particularly true when we encounter non-trivial levels of noise as is the case for S&P 500 index options. In contrast, for other asset classes, such as over the counter currency options which are actually quoted in terms of implied volatilities, virtually any common method of risk neutral density estimation leads to plausible results.[16]

**X Conclusion**

The Black-Scholes formula and its subtle variants are a powerful set of tools useful for the purpose of pricing and hedging options in practice. In this paper I present a method which combines the Black Scholes formula with local polynomial regression to accommodate a non-constant time-evolving volatility smile. My method proves useful for pricing options, estimating option deltas and gammas, and as a means of estimating the risk neutral distribution of the underlying asset at expiration. The advantages of using my approach over other methods of risk neutral density estimation has also been discussed. Given the general nature of statistical analyses where more data is usually considered better, both researcher and practitioner alike can take comfort in a method which fully exploits all available information contained in a single day's set of option quotes. While taking advantage of a richer data set, many of the common problems encountered with high frequency data such as asynchronous trading, can be dealt with in a direct fashion by making use of "target times" for estimating the functions of interest. Also, when the risk neutral density estimates from this model are matched to parametric edgeworth expansion type densities, we virtually always obtain a plausible risk neutral density.

As a final note, the use of local polynomial regression to smooth various financial phenomena is not restricted to option pricing. Given the blessing of typically large intra-day data sets, the field of finance is well-suited to using data-intensive methods such as those described in this paper for other applications as well.

---

[16]See for example Campa, Chang and Reider (1997) for a discussion of alternative risk neutral density estimators as applied to currencies..

## References


Abken, P.A., 1995, "Using Eurodollar Futures Options: Gauging the Market's View of Interest Rate Movements," *Economic Review*, Federal Reserve Bank of Atlanta, pp. 10-30.

Ait-Sahalia, Y. and A. Lo. "Nonparametric Estimation of State-Price Densities Implicit in Financial Asset Prices", working paper, 1995.

Bates, D., 1996, "Jumps and Stochastic Volatility: Exchange Rate Processes Implicit in Deutsche Mark Options." *Review of Financial Studies*, vol 9, No 1, 69-107.

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

Breeden, D. and R. Litzenberger, 1977, "Prices of State-Contingent Claims Implicit in Option Prices," *Journal of Business*, vol. 51, pp. 621-651.

Campa, J., K. Chang and R Reider, 1997, "Implied Exchange Rate Distributions: Evidence from OTC Option Markets", Working Paper.

Chriss, N. *Black Scholes and Beyond*, Chicago: Irwin Professional Publishing, 1997.

Derman, Emanuel and I. Kani, 1994, "Riding on a Smile", *Risk*, Vol. 7, No. 2.

Dupire, Bruno, "Pricing with a Smile", *Risk*, Vol 7, No. 1

Hardle, W. *Applied Nonparametric Regression*. Cambridge: Cambridge University Press, 1989.

Heston, S., 1993, "A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options", *Review of Financial Studies*, Vol. 6, No. 2, pp. 327-343.

Hull, J. *Options, Futures and Other Derivative Securities*. 2nd e. Englewood Cliffs: Prentice Hall, 1993.

Jackwerth, J. C., and M. Rubinstein, 1996, "Recovering Probability Distributions from Contemporaneous Security Prices." *Journal of Finance*, 51, 5, 1611-1631.

Johnson, N.L., and S. Kotz, 1970, *Continuous Univariate Distributions - 1 and 2*, John Wiley & Sons, New York.

Lamoureux, C. G., and W.D. Lastrapes, 1993, "Forecasting Stock-Return Variance: Toward an Understanding of Stochastic Implied Volatilities," *Review of Financial Studies*, 6, 293-326.

Longstaff, F, "Option Pricing and the Martingale Restriciton", *Review of Financial Studies*, 8, 1091-1124.

Malz, A., 1996a, "Using Option Prices to Estimate Realignment Probabilities in the European Monetary System: The Case of Sterling-Mark, *Journal of International Money and Finance*, vol 15, pp. 717-748.

Malz, A., 1996b, "Option-Based Estimates of the Probability Distribution of Exchange Rates and Currency Excess Returns", Federal Reserve Bank of New York, Working paper.

Masson, J. and S. Perrakis, 1997, "A Jumping Smile", working paper.

McCauley, R. and W. Melick, 1996a, "Risk Reversal risk", *Risk*, vol. 9., no. 11, pp. 54-57.

McCauley, R. and W. Melick, 1996b, "Propensity and Density", *Risk*, vol. 9., no. 12, pp. 52-54.

Melick, W. and C. Thomas, 1997, "Using Option Prices to Infer Pdfs for Asset Prices: An Application to Oil Prices During the Gulf Crisis," *Journal of Financial and Quantitative Analysis*, forthcoming.

Rubinstein, M., 1994, "Implied Binomial Trees" *Journal of Finance*, 49, 3, 771-818.

Ruppert, D. (1997). Empirical-bias Bandwidths for Local Polynomial Nonparametric Regression and Density Estimation, *JASA*, 92, 1049-1062.

Ruppert, D., Wand, M.P., Holst, U., and Hössjer, O. (1997). Local Polynomial Variance Function Estimation, *Technometrics*, 39, 262--273.

Ruppert, D. (1997). Local polynomial regression and its applications in environmental statistics, In *Statistics for the Environment, Volume 3* (V. Barnett and F. Turkman, editors) Chicester: John Wiley.

Shimko, D., 1993, "Bounds of Probability" *Risk*, 6, 33-37.

Stein, M. and J. Stein, 1991, "Stock Price Distributions with Stochastic Volatility: An Analytic Approach", *Review of Financial Studies*, Vol. 4, No. 4.

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

## Appendix A

**Calculations for Δ,γ and f(s)**

In this appendix we show how to calculate option deltas, gammas as well the risk neutral pdf of the index at expiration, using the semiparametric implied volatility model given below:

$$\hat{c}_{it} = f(M_{it},\hat{\sigma}(M_{it},t)) = \Phi(d_1) - \frac{e^{(-r\tau)}\Phi(d_2)}{M_{it}}$$

$$d_1 = \frac{\ln(M_{it}) + (r + \hat{\sigma}(M_{it},t)^2/2)\tau}{\hat{\sigma}(M_{it},t)\sqrt{\tau}}$$

$$d_2 = d_1 - \hat{\sigma}(M_{it},t)\sqrt{\tau}$$

Recall that small case "c" refers to call prices which have been standardized by dividing by the dividend adjusted index level. Therefore actual fitted call prices are given as follows:

$$\hat{C}_{it} = (S_t - Divs)\cdot\hat{c}_{it} = (S_t - Divs)\cdot f(M_{it},\hat{\sigma}(M_{it},t))$$

Where ($S_t$-Divs) is a known value, and using local polynomial smoothing methods we estimate the following for the target point ($M_{it}$,t):

$$V_{it} = \hat{\sigma}(M_{it},t) \quad , \quad V_{it}^{/} = \frac{dV_{it}}{dM_{it}} \quad , \quad V_{it}^{//} = \frac{d^2V_{it}}{dM_{it}^2}$$

We now show how to calculate the functions of interest, starting from the most general formulas and proceeding into more and more detail. For convenience we drop the I and t subscripts throughout. Also recall that f(.) refers to the standardized Black-Scholes formula, while $\tilde{f}(.)$ refers to the risk neutral pdf.

$$C = (S - Divs)\cdot f(M,\sigma(M,t)) = \tilde{S}\cdot f(M,\sigma(M,t)) \quad , \; where \quad \tilde{S} = S - Divs$$

$$\Delta = \frac{dC}{dS} = \frac{dC}{d\tilde{S}} = f(M,\sigma(M,t)) + \tilde{S}\frac{df}{d\tilde{S}}$$

$$\gamma = \frac{d\Delta}{dS} = \frac{d^2C}{dS^2} = \frac{d^2C}{d\tilde{S}^2} = 2\frac{df}{d\tilde{S}} + \tilde{S}\cdot\frac{d^2f}{d\tilde{S}^2}$$

$$\tilde{f}(X) = e^{-r\tau}\left(\frac{d^2C}{dX^2}\right) = e^{-r\tau}\cdot\tilde{S}\cdot\left(\frac{d^2f}{dX^2}\right)$$

Where:

$$\frac{df}{d\tilde{S}} = \left(\frac{df}{dM}\right)\cdot\left(\frac{dM}{d\tilde{S}}\right) = \left(\frac{df}{dM}\right)\cdot\left(\frac{1}{X}\right)$$

$$\frac{d^2f}{d\tilde{S}^2} = \left(\frac{d^2f}{dM^2}\right)\cdot\left(\frac{1}{X}\right)^2$$

$$\frac{d^2f}{dX^2} = \left(\frac{M^2}{X^2}\right)\left(\frac{d^2f}{dM^2}\right) + 2\cdot\left(\frac{df}{dM}\right)\left(\frac{M}{X^2}\right)$$

Now solving for $\frac{df}{dM}$ *and* $\frac{d^2f}{dM^2}$ :

$$\frac{df}{dM} = \phi(d_1)\frac{dd_1}{dM} - \frac{e^{-r\tau}\phi(d_2)}{M}\frac{dd_2}{dM} + \frac{e^{-r\tau}\Phi(d_2)}{M^2}$$

$$\textit{where:}\quad \frac{dd_1}{dM} = \frac{\partial d_1}{\partial M} + \frac{\partial d_1}{\partial \sigma}\cdot\frac{\partial \sigma}{\partial M}$$

$$\frac{dd_2}{dM} = \frac{\partial d_2}{\partial M} + \frac{\partial d_2}{\partial \sigma}\cdot\frac{\partial \sigma}{\partial M}$$

$$\frac{\partial d_1}{\partial M} = \frac{\partial d_2}{\partial M} = \frac{1}{M\sigma(M,t)\sqrt{\tau}}$$

$$\frac{\partial d_1}{\partial \sigma} = -\frac{\ln(M)+r\tau}{\sigma(M,t)^2\sqrt{\tau}} + \frac{\sqrt{\tau}}{2}$$

$$\frac{\partial d_2}{\partial \sigma} = -\frac{\ln(M)+r\tau}{\sigma(M,t)^2\sqrt{\tau}} - \frac{\sqrt{\tau}}{2}$$

$$\frac{d^2f}{dM^2} = \phi(d_1)\cdot\left(\frac{d^2d_1}{dM^2}\right) - d_1\phi(d_1)\left(\frac{dd_1}{dM}\right)^2$$

$$- \frac{e^{-r\tau}\phi(d_2)}{M}\cdot\left(\frac{d^2d_2}{dM^2}\right) + \left(\frac{e^{-r\tau}\phi(d_2)}{M^2}\right)\cdot\left(\frac{dd_2}{dM}\right)$$

$$+ \left(\frac{e^{-r\tau}d_2\phi(d_2)}{M}\right)\cdot\left(\frac{dd_2}{dM}\right)^2 - 2\left(\frac{e^{-r\tau}\Phi(d_2)}{M^3}\right)$$

$$+ \left(\frac{e^{-r\tau}\phi(d_2)}{M^2}\right)\cdot\left(\frac{dd_2}{dM}\right)$$

$$\frac{d^2 d_1}{dM^2} = -\frac{1}{M^2 \sigma(M,t)\sqrt{\tau}} - \frac{1}{M\sigma(M,t)^2\sqrt{\tau}} \cdot V' +$$

$$V''\left(\frac{\sqrt{\tau}}{2} - \frac{\ln(M) + r\tau}{\sigma(m,t)^2\sqrt{\tau}}\right) + V'\left[2V'\left(\frac{\ln(M) - r\tau}{\sigma(m,t)^3\sqrt{\tau}}\right) - \frac{1}{M\sigma(M,t)^2\sqrt{\tau}}\right]$$

$$\frac{d^2 d_2}{dM^2} = -\frac{1}{M^2 \sigma(M,t)\sqrt{\tau}} - \frac{1}{M\sigma(M,t)^2\sqrt{\tau}} \cdot V' +$$

$$V''\left(-\frac{\sqrt{\tau}}{2} - \frac{\ln(M) + r\tau}{\sigma(m,t)^2\sqrt{\tau}}\right) + V'\left[2V'\left(\frac{\ln(M) - r\tau}{\sigma(m,t)^3\sqrt{\tau}}\right) - \frac{1}{M\sigma(M,t)^2\sqrt{\tau}}\right]$$

## Appendix B

### Obtaining Implied Volatilities for S&P 500 Index Options

In this appendix we discuss the choice of S&P 500 Index Options for analysis and the steps necessary to arrive at a set of implied volatilities for a given trading day and expiration. This class of derivatives has receiving particular attention in the literature as S&P 500 index options possess many desirable properties for conducting empirical research given the original assumptions underlying the standard Black Scholes equations. Although far from perfect, index options appear to approximate the Black-Scholes assumptions better than most financial derivatives in existence today. In particular S&P 500 index options are very liquid, have European expirations, have predictable and relatively continuous dividend streams, and the underlying asset (the index) is less likely to exhibit large discontinuous jumps versus other types of traded assets such as individual equities.[17] For these reasons an S&P 500 index options database is constructed using intraday data from the Berkeley Options Database (BODB) in conjunction with other data sources.

The actual steps taken to arrive at implied volatilities are similar in nature to those carried out by Jackwerth and Rubinstein (1996), with certain modifications. The raw data originates from four different sources: the Berkeley Options Data Base (BODB), t-bill rates from the Wall Street Journal, the Standard and Poors Daily Dividend Record and the Chicago Mercantile Exchange's S&P 500 Futures tick-by-tick data. From this raw data the implied volatilities corresponding to the bid-ask midpoint of every single option quote are obtained. The steps necessary to arrive at these implied volatilities are described in this appendix in the sequential order in which they are carried out.

Since there is evidence that the reported values for the S&P 500 index tends to lag the actual level by an average of 5-7 minutes and since traders generally hedge their positions using S&P 500 futures contracts, instead of using the reported index value from the Berkeley tape, index levels were implied from the corresponding futures data using futures spot parity.[18] To arrive at the index implied by futures' contracts, the expected dividend payments from the current trading date to futures settlement date are assumed to be equal to the actual dividends paid.[19] These

---

[17]The predictability of dividend streams was confirmed by running a simple regression of actual dividends paid for each day, on actual dividends paid the year prior (lining up calender dates as close as possible). The R-squared from this regression was 28.9%. When regressing weekly dividend payments on dividend payments from 52 weeks prior we obtain an R-squared coefficient of 71.1%.

[18]See Jackwerth & Rubinstein, Footnote 6.

[19]In practice it is necessary to forecast the actual dividends paid. We side-step the issue of subjectively selecting a particular forecasting technique by using actual dividends paid.

streams of dividend payments are discounted back to current dates using a t-bill rate with maturity as close as possible to the corresponding settlement date. Also, instead of using t-bill rates directly for futures-spot parity, individual repo rates are estimated for each day and each futures settlement date; this is accomplished by first "dumping" the reported S&P 500 index level off the Berkeley tape and lining it up with the futures data on a minute by minute basis. For each settlement date a vector of repo rates is estimated from futures-spot parity corresponding to each minute:

$$F_o - S_o(1+r)^{\tau} + Divs(1+r)^{\tau} = 0$$

$$r = \left(\frac{F_0}{S_0 - Divs}\right)^{\frac{1}{\tau}} - 1 \qquad \textbf{(23)}$$

Where $F_o$ and $S_o$ represent the current futures and spot price, $\tau$ represents time to settlement in years, and Divs represents the present value of actual dividends paid on the index from now until settlement. A single repo rate is then obtained for each settlement date by taking the median of the minute by minute repo rates across the day. With a single repo rate and the present value of dividend payments in hand for each settlement date, the reported index levels from the Berkeley tape are then discarded and implied index levels for each minute are estimated as:

$$S_o = \frac{F_o}{(1+r)^{\tau}} + Divs \qquad \textbf{(24)}$$

Note that this yields a set of three or four implied index levels for each trading minute depending upon the number of unique futures contracts in existence at that date. In order to arrive at a single implied index level for each minute, the median of these 3 or 4 implied indices was then taken. It should be stressed that the reported index levels from the Berkeley tape are only used to estimate a single repo rate for each settlement date; these repo rates were then used to back out implied index levels. Empirically, this method produces an implied index level with a mean value across the entire day which is roughly equal to the mean value of the actual reported index levels on the Berkeley tape. At the same time the **minute by minute values** of the implied index should be more reflective of the index level used to actually price options throughout the day.

After calculating implied index levels using the futures data, implied borrowing and lending rates are calculated using the Berkeley options database based upon put-call parity. For each trading day and unique expiration, all quotes for put and call options with the same strike price and time to expiration are lined up minute by minute. Vectors of implied borrowing and lending rates ($r_b$ and $r_l$) are then calculated from put call parity as:

$$P + S_0 = C + X(1+r)^{-\tau} + Divs$$

$$r_b = \left(\frac{X}{P_B + S_0 - Divs - C_A}\right)^{\frac{1}{\tau}} - 1 \qquad \textbf{(25)}$$

$$r_l = \left(\frac{X}{P_A + S_0 - Divs - C_B}\right)^{\frac{1}{\tau}} - 1$$

Where So is the implied index level, $P_B$ and $P_A$ are the bid and ask prices for a put option with strike X, $C_B$ and $C_A$ are the bid and ask prices for a call option with same strike X, and Divs is the present value of actual dividends paid, where dividends are discounted using the appropriate t-bill rate. To arrive at a single borrowing and lending rate for each day and expiration, the median of each vector is then taken.

Given values for the implied index, implied interest rates and expected dividend payments, implied volatilities for each option quote can then be calculated by inverting the dividend-adjusted Black Scholes formulas for European calls and puts. To do this, volatility parameters are numerically solved for by equating the Black-Scholes theoretical values with the mid-point of each bid and ask quote.[20] This procedure typically yields several thousand implied volatility values for each trading day and expiration class, which are then used as inputs for the model described in the main body of the text.

[20]For purposes of calculating implied volatilities the mean of the lending and borrowing rate was used in conjunction with other fundamentals.

## Appendix C

### Overview of EBBS Approach

In this appendix we briefly summarize the empirical bias bandwidth approach of Ruppert (1997), and describe the various "tuning" parameters which must be set in order to estimate optimal local bandwidths and the variance of the function of interest.

The basic approach of the EBBS method is to separately estimate the bias and variance of each function on a local basis. This is done using several candidate bandwidth values, and the bandwidth which yields the minimum estimated mean squared error at each point is then chosen in the final estimation. These optimal bandwidths are then smoothed across the independent variables so as to provide a reasonable and smooth local bandwidth for estimation. The steps of this procedure are summarized below. Since the method requires choosing several tuning parameter, the parameters which were actually chosen for the purpose of this paper have been placed in { } brackets as they are described.

**Step 1)**

Standardize the data so that the mean and variance of each independent variable are zero and one respectively. Since bandwidths are essentially a distance measure, these values can now be viewed in terms of the standard deviation of the raw data in each direction. The search for optimal bandwidths is then done on a common standardized bandwidth.

**Step 2)**

After standardizing the data, choose a relatively low bandwidth value {h=0.5} and perform a local polynomial regression on the data using this global bandwidth. From this initial estimate of the mean function save the residuals($e_{it}$) and a degrees of freedom correction measure ($\Delta_i$) which are given as follows:

$$\begin{aligned} e_i &= Y_i - \hat{m}(X_i,h) \\ \hat{m}_i(X_i,h) &= S_1 Y_i \\ \Delta_i &= diag(S_1 S_1^{/} - 2S_1) \end{aligned} \qquad \textbf{(26)}$$

where $Y_i$ is the dependent variable, m-hat is a fitted value and $S_1$ is the smoothing matrix used for the initial estimate. The smoothing matrix is a TxT matrix such that the vector of fitted values is given as the product of this matrix and the vector of dependent values ($S_1 y$). In this particular case the i'th row of $S_1$ corresponds to $e'(X_i'W_iX_i)^{-1}X_i'W_iY$ as in Equation 7, with the i'th observation point as the target point, and the initial chosen global bandwidth. $\Delta_i$ is a degrees of

freedom type correction used for variance estimation in the next step.

**Step 3)**

If we smooth the squared residuals from Step 2 with the use of a second smoothing matrix $S_2$, and if we assume m-hat is an unbiased estimate of the mean in step 2, then it follows that:

$$E(S_2 e_i^2) = \sigma^2(1 + S_2\Delta) \quad \textbf{(27)}$$

This motivates the following unbiased estimator for var($\epsilon_i^2$):

$$\hat{\sigma}_i^2(X_i,h) = \left(\frac{S_2 e_i^2}{1+S_2\Delta}\right) \quad \textbf{(28)}$$

Therefore, having obtained a set of residuals from the initial smooth, the variance function of the error term can be estimated by smoothing a set of adjusted squared residuals. In order to do this though it is necessary to choose an optimal local bandwidth for $S_2$. To do this we apply the EBBS method to variance estimation as well, and empirically estimate the mean squared error of the variance estimate as a function of the local bandwidth. Instead of attempting to separately estimate the variance of the variance estimate, it is assumed that the $\epsilon_i$'s come from a scale family so that for some >0:

$$Var(\sigma^2(X_i)\epsilon_i^2) = Var(\epsilon_i^2)\sigma^4(X_i) = \sigma^2(X_i) \quad \textbf{(29)}$$

which suggests the following estimator for the variance of the squared residuals:

$$var(\hat{\sigma}^2(X_i,h)) = \hat{}\hat{\sigma}^4(X_i,h) \quad \textbf{(30)}$$

Since E($e_i$)=0, kappa-hat represents an estimate of the residual kurtosis minus one and can be estimated by partitioning the data into M blocks with $N_l$ (l=1,2,..M) observations in each block, and calculating the average ratio of the sample variance of squared residuals to the sample mean of squared residuals in each data block as follows:

$$k_l = \frac{((\sum_{i=1}^{N_l}(e_i^4)) - (\sum_{i-1}^{N_l} e_i^2)^2)/(N_l - 1)}{\frac{1}{N_i}(\sum_{i=1}^{N_l} e_i^2)} \qquad l=1,2,..M$$

$$\hat{} = \frac{1}{M}(\sum_{l=1}^{M} k_l) \qquad \textbf{(31)}$$

It is therefore possible to estimate the variance function at several target points and using kappa-hat from above, in conjunction with (29), we can obtain the variance of these estimates conditional on the data and each candidate bandwidth. {In this paper, 400 target points were chosen using a 20X20 Cartesian grid linearly spaced in each direction starting at the minimum of each variable and extending to the maximum. Thirty-five geometrically spaced candidate bandwidths were used, ranging from 0.25 to 2. For estimating kappa-hat the data was divided into M=9 groups, based upon a 3x3 grid, } To calculate the corresponding mean squared error, it is also necessary to estimate the bias of the variance estimate associated with each bandwidth; this is described in the following step.

**Step 5)**

The idea behind bias estimation is to fit a model for the expectation of the object of interest as a function of various candidate bandwidths. Letting $x_i$ represent a target point of interest, and $h_0$ a candidate bandwidth for smoothing the function, the bias at a particular point is given approximately as follows:

$$\hat{f}(x_i;h) \approx c_0(x_i) + c_{p+1-K}(x_i)\, h^{p+1-K} + \ldots + c_{p+t-K}\, h^{p+t-K} \qquad \textbf{(32)}$$

f-hat in this case can refer to the mean function, variance function or a derivative of interest. The terms to the right of $c_o(x_i)$ represent an estimate of bias; and the bias coefficients (c's) are estimated by ordinary least squares. In this case fitted values of $f(x_i,h)$ are obtained for several geometrically spaced bandwidths in a neighborhood of $h_0$. These fitted values are then regressed on a polynomial in h which contains t terms plus a constant. p in the above equation continues to represent the order of the local polynomial regression used to fit the data while K represents the order of the function being estimated; K=0 when estimating implied volatility surfaces and the variance of functions, while K=1 and K=2 for first and second derivative estimation. By regressing fitted values on candidate bandwidths in a neighborhood of $h_{o,}$ the estimate of bias at a point of interest is then given as:

$$bias(\hat{f}(x_i;h_o)) = \hat{c}_{p+1-K}(x_i)h_0^{p+1-K} + \ldots. + \hat{c}_{p+t-K}h_o^{p+t-K} \qquad \textbf{(33)}$$

{t is chosen to be equal to 3 for all functions of interest. For estimating the bias at $x_i$ with $h_o$, $J_1$=1 points to the left of $h_o$, and $J_2$=3 points to the right of $h_o$ are used in the fit}. Based on (29), (30) and (32) it is therefore possible to calculate the MSE of a local variance estimate for the candidate bandwidths. To calculate the MSE of the estimated mean function or derivative estimates, we appeal to standard generalized least squares theory. Since local polynomial regression involves a weighted least squares regression at each target point, it follows from the first line of (7) that:

$$\hat{\beta}_i = (X'WX)^{-1}X'WY$$

$$Var(\hat{\beta}_i) = (X'WX)^{-1}(X'W \quad W'X)(X'WX)^{-1}$$

$$Var(\hat{\beta}_i) = \sigma^2(x_i)[(X'WX)^{-1}(X'WW'X)(X'WX)^{-1}] \qquad \textbf{(34)}$$

$$\overset{\wedge}{Var}(\hat{f}(X_i,h)) = \hat{\sigma}^2(x_i)(K!)^2[(X'WX)^{-1}(X'WW'X)(X'WX)^{-1}]_{(K+1)(K+1)}$$

Once again K refers to the order of the function being estimated, and $[\ ]_{(K+1)(K+1)}$ refers to the (K+1)'th diagonal element of the matrix in square brackets. Combining (33) with (32) it is therefore possible to estimate the mean squared error of the surface being estimated and any derivatives of interest. Also, if a common bandwidth is used for estimating the function as well as all derivatives of interest, it is straight forward to obtain the variance-covariance matrix all of these functions based on the non-diagonal elements of (34) above.

**Step 6)**

Estimates of the MSE as described in the steps above are typically quite rough when graphed as a function of the grid of target points for any particular candidate bandwidth. Therefore as a precursor to choosing optimal bandwidths, it is desirable to smooth the MSE function with respect to both variables of interest for every candidate bandwidth considered. {This auxiliary smooth is conducted using local linear fitting, with a global bandwidth set equal to 1.125 - the midpoint of the smallest and largest candidate bandwidths} Having smoothed these functions, optimal bandwidths are then chosen at each target point so as to correspond to the first local minimum of smoothed MSE. A first local minimum is used rather than a global minimum as very large values of h tend to greatly underestimate bias when all major features of f(.) have been completely smoothed away. In fact a global minimum of *estimated* MSE would be reached if h were allowed to tend towards infinity.

**Step 7)**

The very last step in obtaining a set of optimal bandwidths is to conduct a final smooth on the optimal bandwidths corresponding to the first local minimum of smoothed MSE in Step 6 above. {This is also done here, using a global bandwidth set equal to 1.125}

By following the steps above, which comprise the EBBS algorithm, it is possible to obtain smooth yet locally responsive bandwidths, useful for estimating multivariate functions, derivatives of functions, as well as the variance of any function of interest. Out of necessity of space, many of the details of this algorithm have been omitted, and the reader is referred to Ruppert (1997), as well as Ruppert, Wand and Hossjer (1997).

Figure 1—A
Local Linear Fit of I.V. Surface ($h_m|h_t$=0.025|12.5)

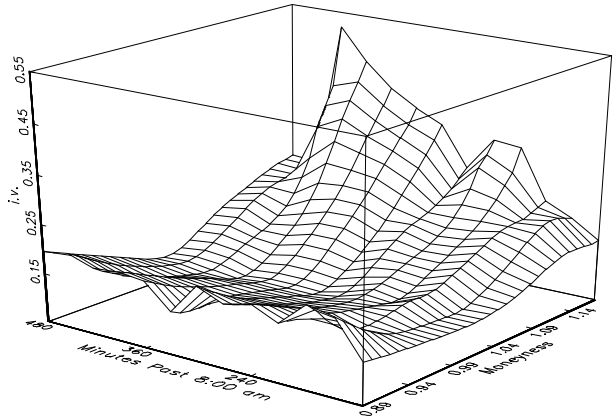



Figure 1—B
Local Linear Fit of I.V. Surface ($h_m|h_t$=0.05|25.0)

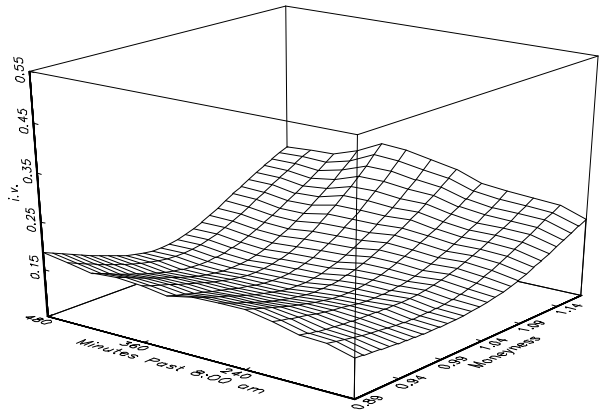



Figure 1—C
Local Linear Fit of I.V. Surface ($h_m|h_t$=0.1|50.0)

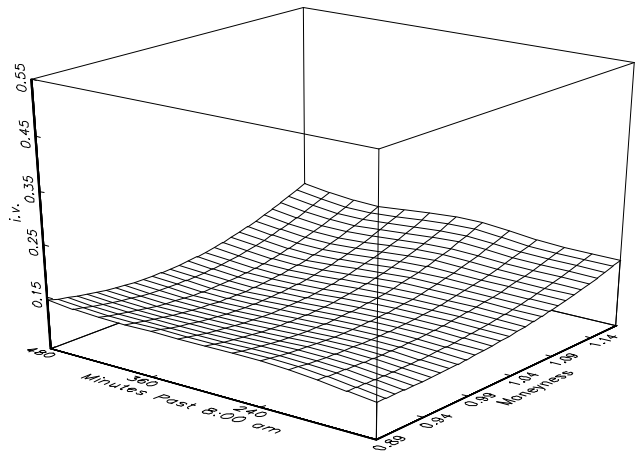



Figure 1:

Impact of a higher bandwidth on the level of smoothing. Moving from top to bottom, more and more smoothing is introduced in an estimate of the implied volatility surface corresponding to January 1, 1997, (March Expiration) using local linear fitting and a common global bandwidth.

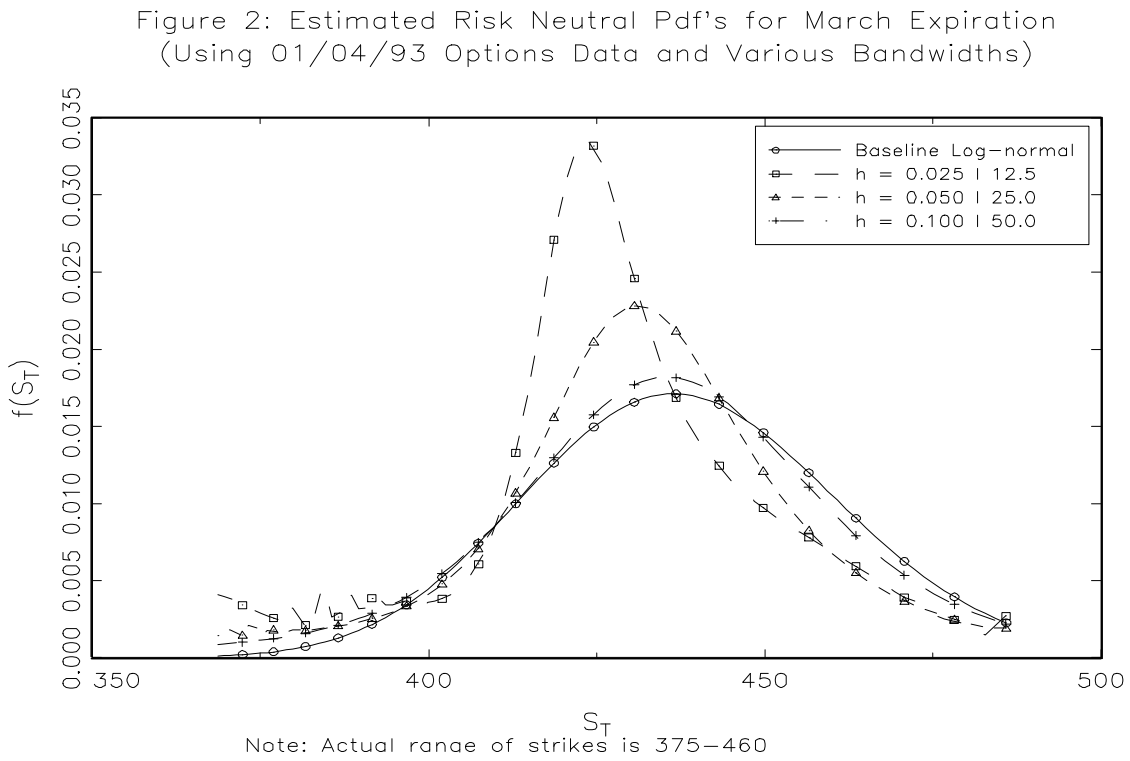



Figure 2:

Risk neutral pdfs corresponding to the implied volatility surfaces of Figure 1. Also included is a base-line log-normal distribution using the median of all implied volatility values for the day. As we introduce more and more smoothing the tails of the estimated distribution become better behaved and the estimated distribution is forced towards the base-line log-normal. Since the surfaces of Figure 1 are estimated using local linear fitting, the second derivative with respect to the strike price is obtained via first differencing the first derivative estimate in this case.

Figure 3–A: Estimated Implied Volatility Surface
01/04/93, March Expiration

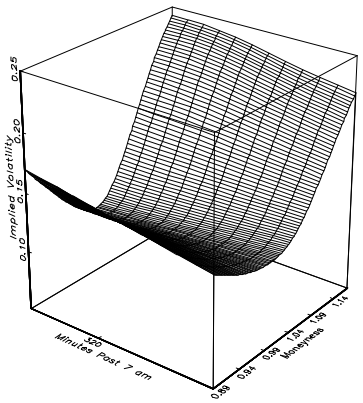



Figure 3–B: Bandwidth for Estimated Implied Volatility Surface
01/04/93, March Expiration

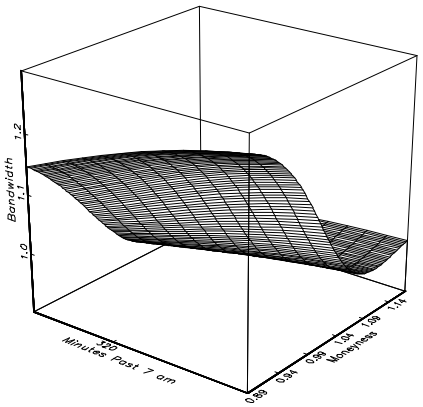



Figure 4–A: Estimated Standard Deviation of Implied Volatility Surface
01/04/93, March Expiration

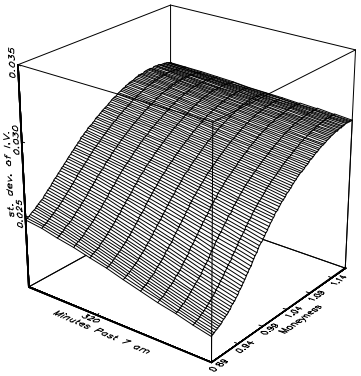



Figure 4–B: Bandwidth for Variance Estimation
01/04/93, March Expiration

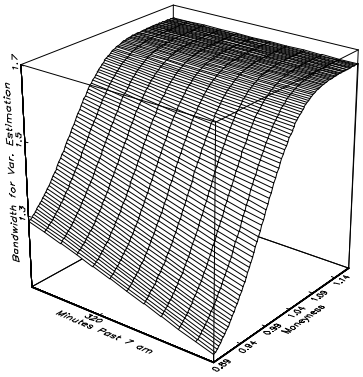



Figure 5–A: Estimated First Derivative wrt Moneyness
01/04/93, March Expiration

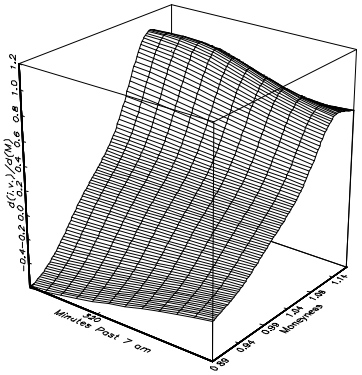
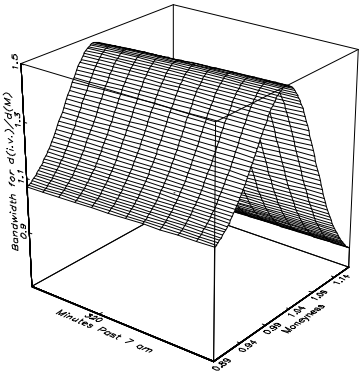



Figure 5–B: Bandwidth for First Derivative Estimation
01/04/93, March Expiration




Figures 3,4,5:
EBBS method in practice. Figures 3A, 4A and 5A illustrate the estimation of 3 functions of interest. Specifically the implied volatility surface, the heteroscedastic variance function and the first derivative of implied volatility with respect to moneyness. To the right of each function is the local bandwidth used for estimation in each case, as chosen by EBBS.

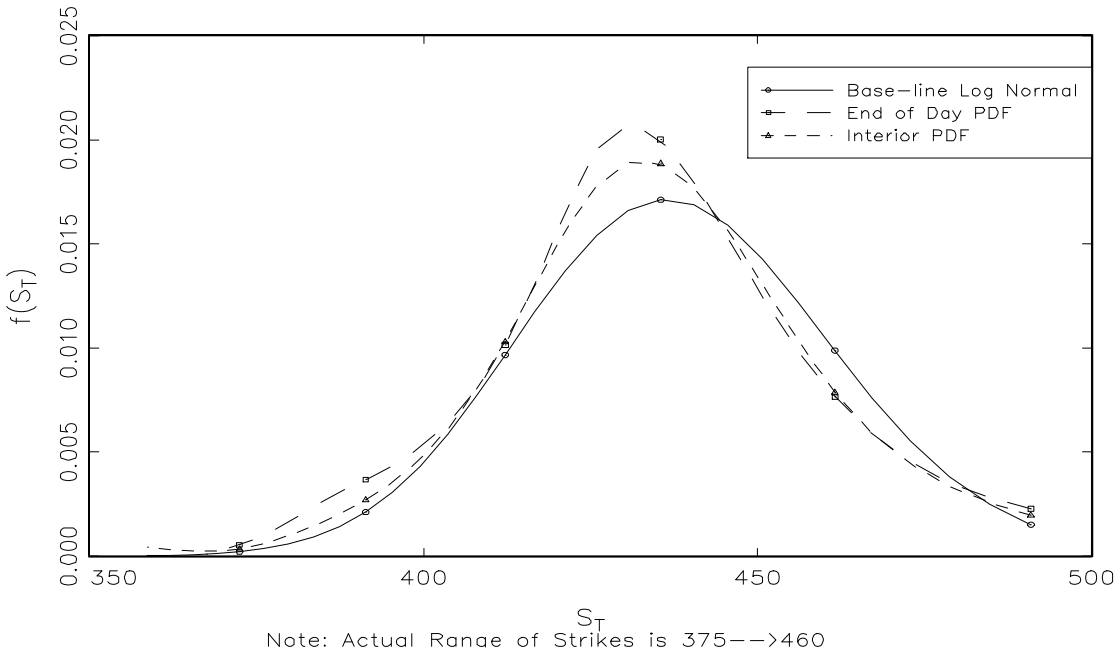



Figure 6:

Estimates of the risk neutral density for January 3, 1993, (March Expiration), corresponding to the surface of 3A, using local linear fitting and empirically driven bandwidths. Two estimates of the risk neutral pdf are provided corresponding to different times of the day.

Figure 7–A: Estimated Implied Volatility Surface using EBBS
01/04/97, January Expiration

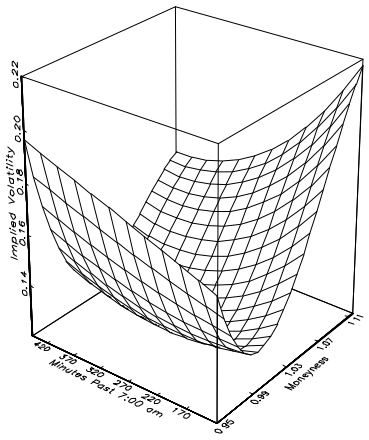



Figure 7–B: Estimated Implied Volatility Surface using EBBS
01/04/97, February Expiration

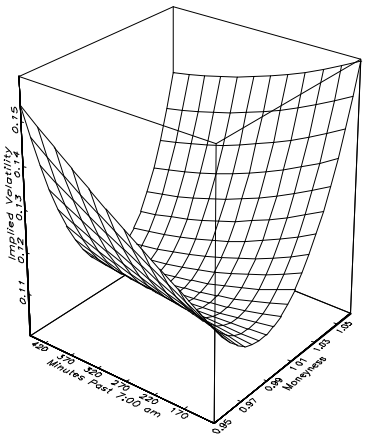



Figure 7–C: Estimated Implied Volatility Surface using EBBS
01/04/97, March Expiration

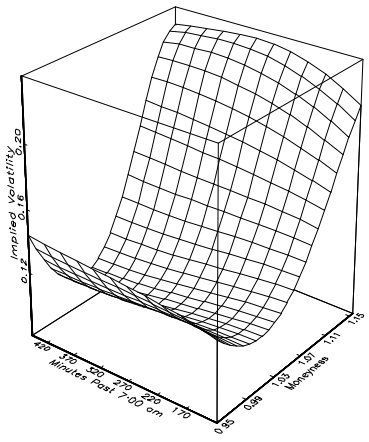



Figure 7–D: Estimated Implied Volatility Surface using EBBS
01/04/93, June Expiration

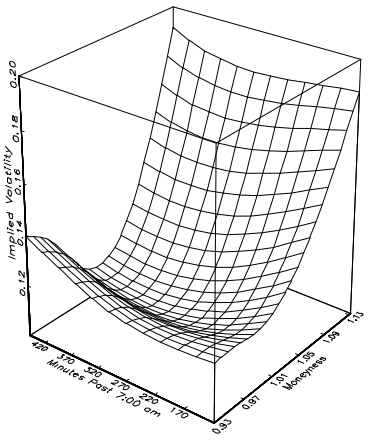
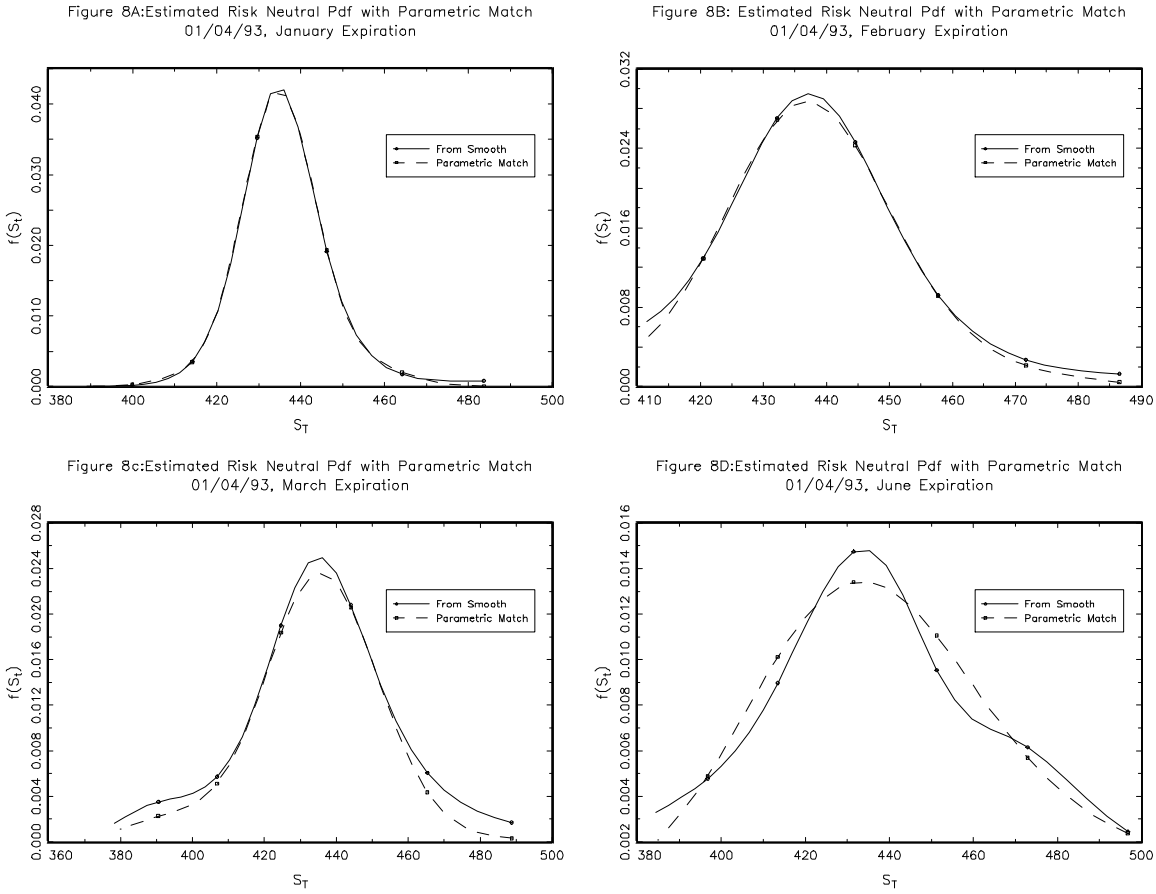



Figure 7:

Some sample implied volatility surfaces for January 4, 1993, corresponding to the first four expiration dates. Note the commonly observed smile where deep in, and deep out of the money options tend to have higher implied volatilities. Also, notice the presence of intra-day changes in the shape and level of the smile, particularly for the closest two expiration dates.

Figure 8:




Estimated Pdfs corresponding to the surfaces of Figure 7. Also included are 4 parameter edgeworth expansion type densities which are chosen so as to minimize the mean integrated square error between the estimated pdf and the parametric match over the range of actually observed strike prices.

Figure 9: Estimated Risk Neutral Pdf's using Parametric Match
01/04/93

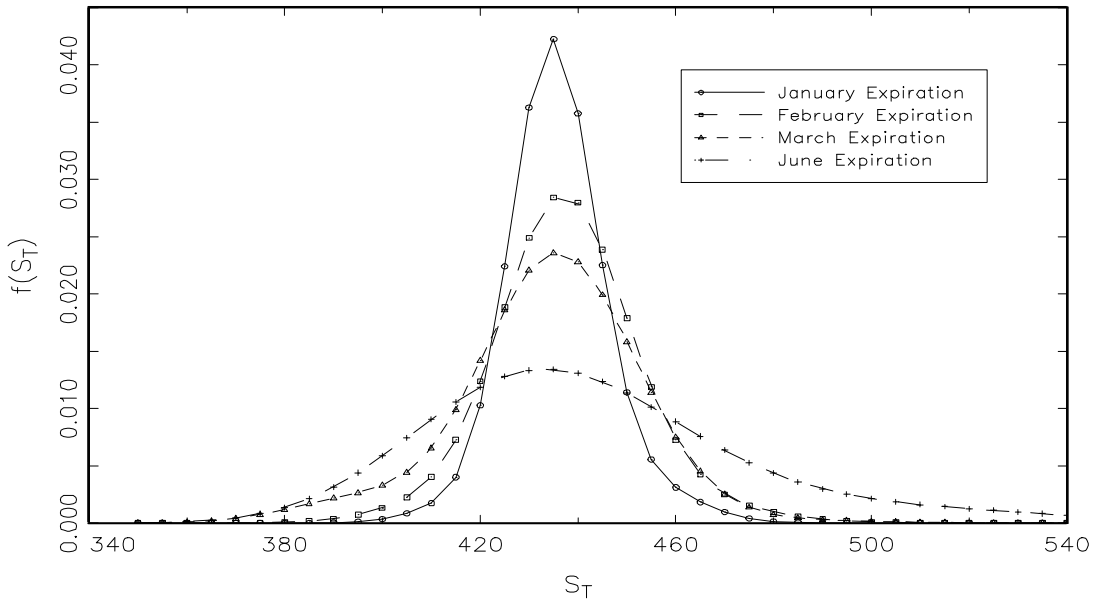



Figure 9:

The parametrically matched pdfs from the four panels of Figure 8 are combined into a single graph so as to demonstrate the increased uncertainty associated with the farther term expiration dates.

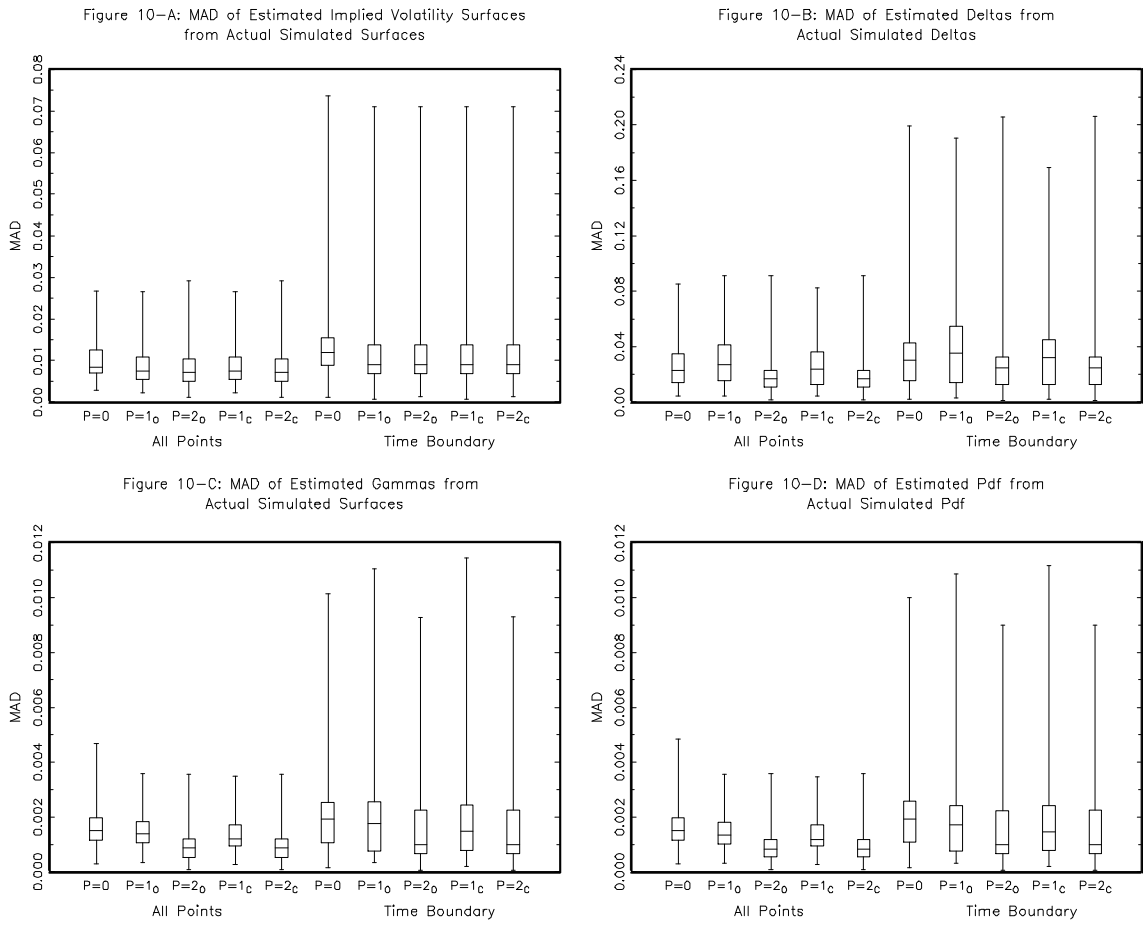



Figure 10:

Performance of the EBBS method at interior points and the time boundary in terms of estimating the various functions of interest. In this case the mean absolute deviation of the estimated functions from their actual simulated values are calculated across all target points in the estimation.

The different box plots correspond to different ordered polynomial schemes. An “O” subscript refers to optimal local bandwidths chosen for all functions, whereas a “C” subscript implies a common optimal local bandwidth is used for estimating the underlying function and derivatives of interest. Simple differencing is used for lower ordered polynomials where necessary, in order to estimate the second derivative of implied volatility with respect to moneyness.

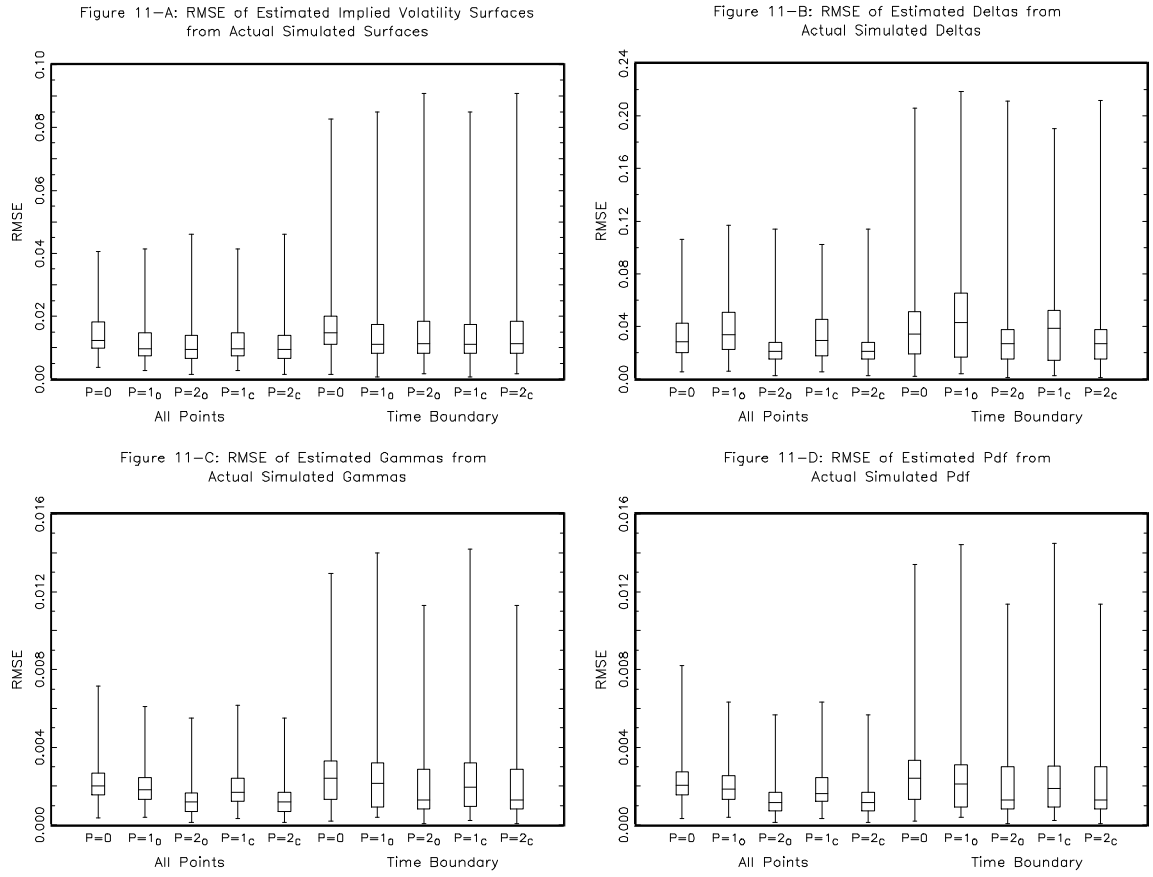



Figure 11:

Same as Figure 10, using a root mean squared error criteria, across all target points.

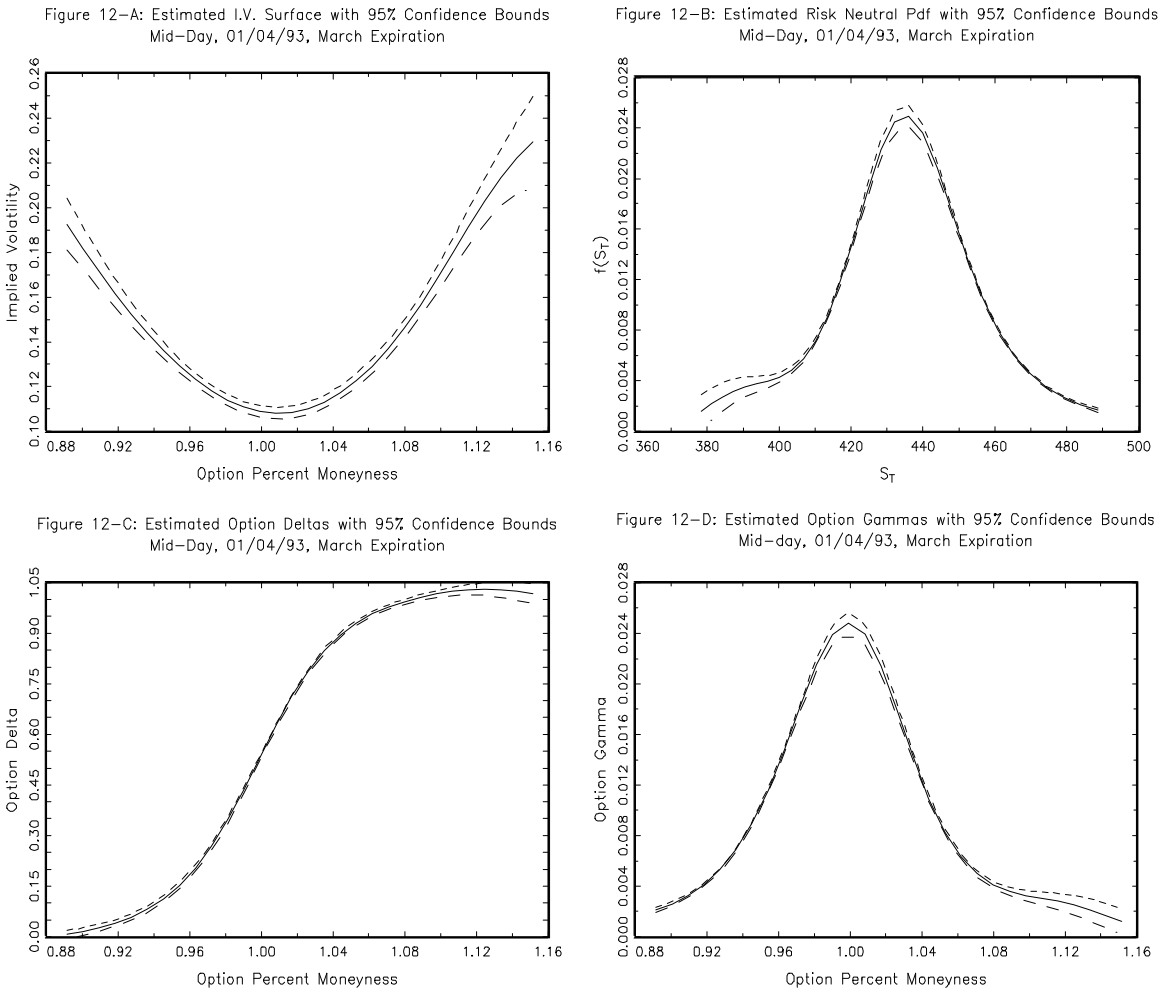



Figure 12:

Estimated functions of interest with 95% confidence bounds. Confidence bounds are obtained by drawing from a multivariate normal distribution with variance covariance matrix equal to the estimated variance covariance matrix for the three underlying functions (implied volatility surface, and first and second derivative with respect to moneyness) at each target point.

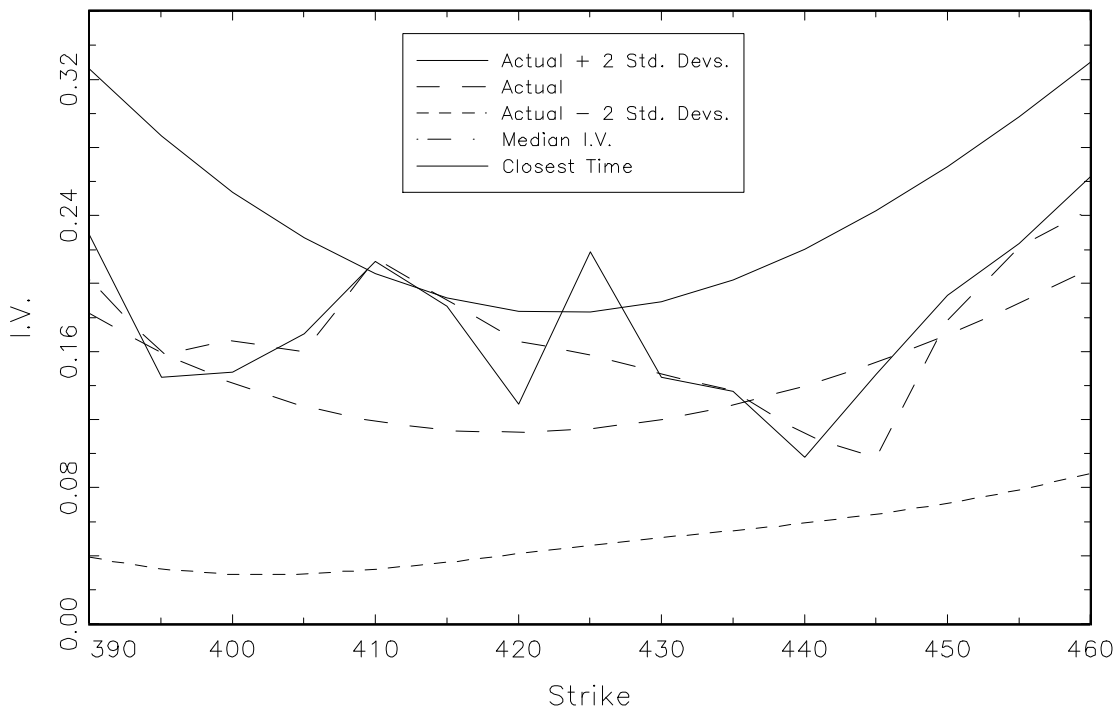



Figure 13:

Implied volatility smile corresponding to first simulation run. Included is the actual smile plus and minus two standard deviations of the data, as well as the smile corresponding to median implied volatility values and closest quote to mid-day implied volatility values for each strike.

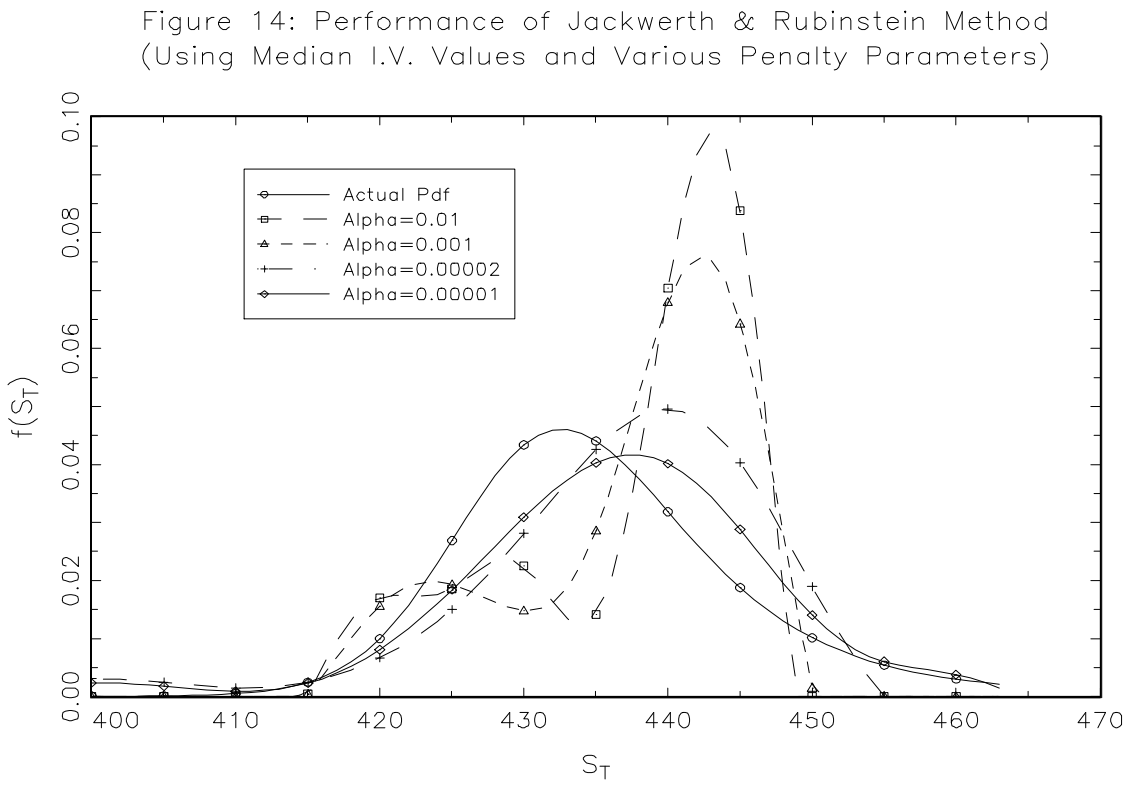



Figure 14:

Performance of Jackwerth and Rubinstein method in recovering actual distribution from median implied volatility values, using various penalty parameters.

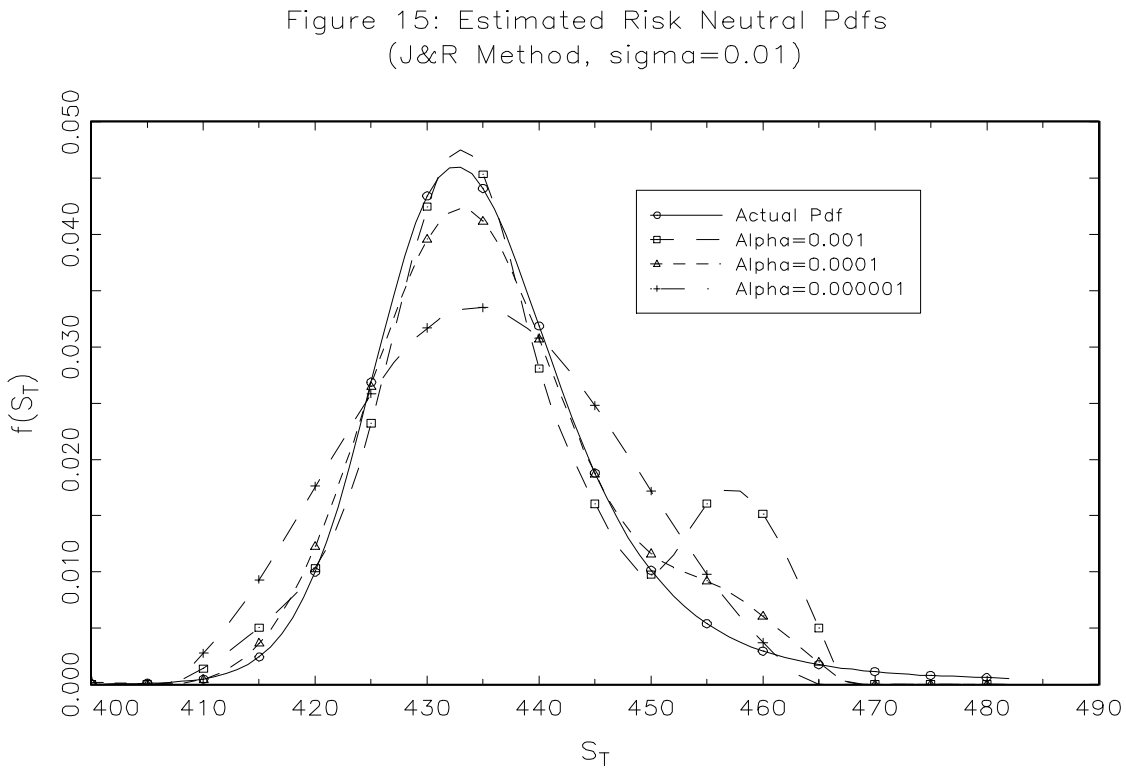



Figure 15:

Performance of Jackwerth and Rubinstein method in recovering actual risk neutral distribution in the face of well-behaved error terms.