Version of July 30, 2008

forthcoming in

Volatility and Time Series Econometrics: Essays in Honor of Robert F. Engle (eds. Tim Bollerslev, Jeffrey R. Russell and Mark Watson). Oxford, UK: Oxford University Press, 2008.

Estimating the Implied Risk Neutral Density for the U.S. Market Portfolio

by

Stephen Figlewski*

* Professor of Finance
New York University Stern School of Business
44 West 4th Street, Suite 9-160
New York, NY 10012-1126

212-998-0712
sfiglews@stern.nyu.edu

Thanks to Justin Birru for excellent assistance on this research and to Otto van Hemert, Robert Bliss, Tim Bollerslev, an anonymous reviewer, and seminar participants at NYU, Baruch, Georgia Tech, Essex University, Lancaster University, Bloomberg, and the Robert Engle Festschrift Conference for valuable comments.

The Implied Risk Neutral Density for the U.S. Market Portfolio

Abstract

The market's risk neutral probability distribution for the value of an asset on a future date can be extracted from the prices of a set of options that mature on that date, but two key technical problems arise. In order to obtain a full well-behaved density, the option market prices must be smoothed and interpolated, and some way must be found to complete the tails beyond the range spanned by the available options. This paper develops an approach that solves both problems, with a combination of smoothing techniques from the literature modified to take account of the market's bid-ask spread, and a new method of completing the density with tails drawn from a Generalized Extreme Value distribution. We extract twelve years of daily risk neutral densities from S&P 500 index options and find that they are quite different from the lognormal densities assumed in the Black-Scholes framework, and that their shapes change in a regular way as the underlying index moves. Our approach is quite general and has the potential to reveal valuable insights about how information and risk preferences are incorporated into prices in many financial markets.

Introduction

The Black-Scholes (BS) option pricing model has had an enormous impact on academic valuation theory and also, impressively, in the financial marketplace. It is safe to say that virtually all serious participants in the options markets are aware of the model and most use it extensively. Academics tend to focus on the BS model as a way to value an option as a function of a handful of intuitive input parameters, but practitioners quickly realized that one required input, the future volatility of the underlying asset, is neither observable directly nor easy to forecast accurately. However, an option's price in the market is observable, so one can invert the model to find the implied volatility (IV) that makes the option's model value consistent with the market. This property is often more useful than the theoretical option price for a trader who needs the model to price less liquid options consistently with those that are actively traded in the market, and to manage his risk exposure.

An immediate problem with IVs is that when they are computed for options written on the same underlying they differ substantially according to "moneyness." The now-familiar pattern is called the volatility smile, or for options on equities, and stock indexes in particular, the smile has become sufficiently asymmetrical over time, with higher IVs for low exercise price options, that it is now more properly called a "smirk" or a "skew."[1]

Implied volatility depends on the valuation model used to extract it, and the existence of a volatility smile in Black-Scholes IVs implies that options market prices are not fully consistent with that model. Even so, the smile is stable enough over short time intervals that traders use the BS model anyway, by inputting different volatilities for different options according to their moneyness. This jury-rigged procedure, known as "practitioner Black-Scholes," is an understandable strategy for traders, who need some way to impose pricing consistency across a broad range of related financial instruments, and don't care particularly about theoretical consistency with academic models. This has led to extensive analysis of the shape and dynamic behavior of volatility smiles, even though it is odd to begin with a model that is visibly inconsistent with the empirical data and hope to improve it by modeling the behavior of the inconsistency.

Extracting important but unobservable parameters from option prices in the market is not limited to implied volatility. More complex models can be calibrated to the market by implying out the necessary parameter values, such as the size and intensity of discrete price jumps. The most fundamental valuation principle, which applies to all financial assets, not just options, is that a security's market price should be the market's expected

[1] Occasionally a writer will describe the pattern as a "sneer" but this is misleading. A smile curves upward more or less symmetrically at both ends; a smirk also curves upward but more so at one end than the other; a "skew" slopes more or less monotonically downward from left to right; but the term "sneer" would imply a downward curvature, i.e., a concave portion of the curve at one end, which is not a pattern seen in actual options markets.

value of its future payoff, discounted back to the present at a discount rate appropriately adjusted for risk.

Risk premia are also unobservable, unfortunately, but a fundamental insight of contingent claims pricing theory is that when a pricing model can be obtained using the principle of no-arbitrage, the risk neutral probability distribution can be used in computing the expected future payoff, and the discount rate to bring that expectation back to the present is the riskless rate. The derivative security can be priced relative to the underlying asset under the risk-neutralized probability distribution because investors' actual risk preferences are embedded in the price of the underlying asset.

Breeden and Litzenberger (1978) and Banz and Miller (1978) showed that, like implied volatility, the entire risk neutral probability distribution can be extracted from market option prices, given a continuum of exercise prices spanning the possible range of future payoffs. An extremely valuable feature of this procedure is that it is model-free, unlike extracting IV. The risk neutral distribution does not depend on any particular pricing model.

At a point in time, the risk-neutral probability distribution and the associated risk neutral density function, for which we will use the acronym RND, contain an enormous amount of information about the market's expectations and risk preferences, and their dynamics can reveal how information releases and events that affect risk attitudes impact the market. Not surprisingly, a considerable amount of previous work has been done to extract and interpret RNDs, using a variety of methods and with a variety of purposes in mind.[2]

Estimation of the RND is hampered by two serious problems. First, the theory calls for options with a continuum of exercise prices, but actual options markets only trade a relatively small number of discrete strikes. This is especially problematic for options on individual stocks, but even index options have strikes at least 5 points apart, and up to 25 points apart or more in some parts of the available range. Market prices also contain microstructure noise from various sources, and bid-ask spreads are quite wide for options, especially for less liquid contracts and those with low prices. Slight irregularities in observed option prices can easily translate into serious irregularities in the implied RND, such as negative probabilities. Extracting a well-behaved estimate of an RND requires interpolation, to fill in option values for a denser set of exercise prices, and smoothing, to reduce the influence of microstructure noise.

The second major problem is that the RND can be extracted only over the range of available strikes, which generally does not extend very far into the tails of the distribution. For some purposes, knowledge of the full RND is not needed. But in many

[2] An important practical application of this concept has been the new version of the Chicago Board Options Exchange's VIX index of implied volatility (Chicago Board Options Exchange (2003)). The original VIX methodology constructed the index as a weighted average of BS implied volatilities from 8 options written on the S&P 100 stock index. This was replaced in 2003 by a calculation that amounts to estimating the standard deviation of the risk neutral density from options on the S&P 500 index.

cases, what makes options particularly useful is the fact that they have large payoffs in the comparatively rare times when the underlying asset makes a large price move, i.e., in the tails of its returns distribution.

The purpose of this paper is to present a new methodology for extracting complete well-behaved RND functions from options market prices and to illustrate the potential of this tool for understanding how expectations and risk preferences are incorporated into prices in the U.S. stock market. We review a variety of techniques for obtaining smooth densities from a set of observed options prices and select one that offers good performance. This procedure is then modified to incorporate the market's bid-ask spread into the estimation. Second, we will show how the tails of the RND obtained from the options market may be extended and completed by appending tails from a Generalized Extreme Value (GEV) distribution. We then apply the procedure to estimate RNDs for the S&P 500 stock index from 1996 - 2008 and develop several interesting results.

The next section will give a brief review of the extensive literature related to this topic. Section 3 details how the RND can theoretically be extracted from options prices. The following section reviews alternative smoothing procedures needed to obtain a well-behaved density from actual options prices. Section 5 presents our new methodology for completing the RND by appending tails from a GEV distribution. Section 6 applies the methodology to explore the behavior of the empirical RND for the Standard and Poor's 500 index over the period 1996-2008. The results presented in this section illustrate some of the great potential of this tool for revealing how the stock market behaves. The final section will offer some concluding comments and a brief description of several potentially fruitful lines of future research based on this technology.

2. Review of the Literature

The literature on extracting and interpreting the risk neutral distribution from market option prices is broad, and it becomes much broader if the field is extended to cover research on implied volatilities and on modeling the returns distribution. In this literature review, we restrict our attention to papers explicitly on RNDs.

The monograph by Jackwerth (2004) provides an excellent and comprehensive review of the literature on this topic, covering both methodological issues and applications. Bliss and Panigirtzoglou (2002) also give a very good review of the alternative approaches to extracting the RND and the problems that arise with different methods. Bahra (1997) is another often-cited review of methodology, done for the Bank of England prior to the most recent work in this area.

One way to categorize the literature is according to the methods used by different authors to extract an RND from a set of option market prices. These fall largely into three approaches: fitting a parametric density function to the market data, approximating the RND with a nonparametric technique, or developing a model of the returns process that produces the empirical RND as the density for the value of the underlying asset on option expiration day.

An alternative classification is according to the authors' purpose in extracting a risk neutral distribution. Many authors begin with a review of the pros and cons of different extraction techniques in order to select the one they expect to work best for their particular application. Because a risk neutral density combines estimates of objective probabilities and risk preferences, a number of papers seek to use the RND as a window on market expectations about the effects of economic events and policy changes on exchange rates, interest rates, and stock prices. Other papers take the opposite tack, in effect, abstracting from the probabilities in order to examine the market's risk preferences that are manifested in the difference between the risk neutral density and the empirical density. A third branch of the literature is mainly concerned with extracting the RND as an econometric problem. These papers seek to optimize the methodology for estimating RNDs from noisy market options prices. The most ambitious papers construct an implied returns process, such as an implied binomial tree, that starts from the underlying asset's current price and generates the implied RND on option expiration date. This approach leads to a full option pricing model, yielding both theoretical option values and Greek letter hedging parameters.

Bates (1991) was one of the first papers concerned with extracting information about market expectations from option prices. It analyzed the skewness of the RND from S&P500 index options around the stock market crash of 1987 as a way to judge whether the crash was anticipated by the market. Like Bahra (1997), Söderlind and Svensson (1997) proposed learning about the market's expectations for short term interest rates, exchange rates, and inflation by fitting RNDs as mixtures of two normal or lognormal densities. Melick and Thomas (1997) modeled the RND as a mixture of three lognormals. Using crude oil options, their estimated RNDs for oil prices during the period of the 1991 Persian Gulf crisis were often bimodal and exhibited shapes that were inconsistent with a univariate lognormal. They interpreted this as the market's response to media commentary at the time and the anticipation that a major disruption in world oil prices was possible.

In their examination of exchange rate expectations, Campa, Chang and Reider (1998) explored several estimation techniques and suggested that there is actually little difference among them. However, this conclusion probably depends strongly on the fact that their currency options data only provided five strike prices per day, which substantially limits the flexibility of the functional forms that could be fitted. Malz (1997) also modeled exchange rate RNDs and added a useful wrinkle. FX option prices are typically quoted in terms of their implied volatilities under the Garman-Kohlhagen model and moneyness is expressed in terms of the option's delta. For example, a "25 delta call" is an out of the money call option with a strike such that the option's delta is 0.25. Malz used a simple function involving the prices of option combination positions to model and interpolate the implied volatility smile in delta-IV space.

Quite a few authors have fitted RNDs to stock market returns, but for the most part, their focus has not been on the market's probability estimates but on risk preferences. An exception is Gemmill and Saflekos (2000), who fitted a mixture of two lognormals to FTSE stock index options and looked for evidence that investors' probability beliefs prior to British elections reflected the dichotomous nature of the possible outcomes.

Papers that seek to use RNDs to examine the market's risk preferences include Ait-Sahalia and Lo (1998, 2000), Jackwerth (2000), Rosenberg and Engle (2002) and Bliss and Panigirtzoglou (2004).

In their 1998 paper, Ait-Sahalia and Lo used a nonparametric kernel smoothing procedure to extract RNDs from S&P 500 index option prices. Unlike other researchers, they assumed that if the RND is properly modeled as a function of moneyness and the other parameters that enter the Black-Scholes model, it will be sufficiently stable over time that a single RND surface defined on log return and days to maturity can be fitted to a whole calendar year (1993). While they did not specifically state that their approach focuses primarily on risk preferences, it is clear that if the RND is this stationary, its shape is not varying in response to the flow of new information entering the market's expectations, beyond what is reflected in the changes in the underlying asset price. Ait-Sahalia and Lo (2000) applies the results of their earlier work to the Value-at-Risk problem and proposes a new VaR concept that includes the market's risk preferences as revealed in the nonparametric RND.

Jackwerth (2000) uses the methodology proposed in Jackwerth and Rubinstein (1996) to fit smooth RNDs to stock prices around the 1987 stock market crash and the period following it. The paper explores the market's risk attitudes, essentially assuming that they are quite stable over time, but subject to substantial regime changes. The resulting risk aversion functions exhibit some anomalies, however, leaving some important open questions.

In their cleverly designed study, Bliss and Panigirtzoglou (2004) assume a utility function of a particular form. Given a level of risk aversion, they can then extract the representative investor's true (subjective) expected probability distribution. They assume the representative investor has rational expectations and find the value of the constant risk aversion parameter that gives the best match between the extracted subjective distribution and the distribution of realized outcomes. By contrast, Rosenberg and Engle (2002) model a fully dynamic risk aversion function by fitting a stochastic volatility model to S&P 500 index returns and extracting the "empirical pricing kernel" on each date from the difference between the estimated empirical distribution and the observed RND.

The literature on implied trees began with three papers written at about the same time. Perhaps the best-known is Rubinstein's (1994) Presidential Address to the American Finance Association, in which he described how to fit Binomial trees that replicate the RNDs extracted from options prices. Rubinstein found some difficulty in fitting a well-behaved left tail for the RND and chose the approach of using a lognormal density as a

Bayesian prior for the RND. Jackwerth (1997) generalized Rubinstein's binomial lattice to produce a better fit, and Rubinstein (1998) suggested a different extension, using an Edgeworth expansion to fit the RND and then constructing a tree consistent with the resulting distribution. Both Dupire (1994) and Derman and Kani (1994) also developed implied tree models at about the same time as Rubinstein. Dupire fit an implied trinomial lattice, while Derman and Kani, like Rubinstein, used options prices to imply out a binomial tree, but they combined multiple maturities to get a tree that simultaneously matched RNDs for different expiration dates. Their approach was extended in Derman and Kani (1998) to allow implied (trinomial) trees that matched both option prices and implied volatilities. Unfortunately, despite the elegance of these techniques, their ability to produce superior option pricing and hedging parameters was called into question by Dumas, Fleming and Whaley (1999) who offered empirical evidence that the implied lattices were no better than "practitioner Black-Scholes".

The most common method to model the RND is to select a known parametric density function, or a mixture of such functions, and fit its parameters by minimizing the discrepancy between the fitted function and the empirical RND. A variety of distributions and objective functions have been investigated and their relatives strengths debated in numerous papers, including those already mentioned. Simply computing an implied volatility using the Black-Scholes equation inherently assumes the risk neutral density for the cumulative return as of expiration is lognormal. Its mean is the riskless rate (with an adjustment for the concavity of the log function) and it has standard deviation consistent with the implied volatility, both properly scaled by the time to expiration.[3] But given the extensive evidence that actual returns distributions are too fat-tailed to be lognormal, research with the lognormal has typically used a mixture of two or more lognormal densities with different parameters.

Yet, using the Black-Scholes equation to smooth and interpolate option values has become a common practice. Shimko (1993) was the first to propose converting option prices into implied volatilities, interpolating and smoothing the curve, typically with a cubic spline or a low-order polynomial, then converting the smoothed IVs back into price space and proceeding with the extraction of an RND from the resulting dense set of option prices. We adopt this approach below, but illustrate the potential pitfall of simply fitting a spline to the IV data: Since a standard cubic spline must pass through all of the original data points, it incorporates all of the noise from the bid-ask spread and other market microstructure frictions into the RND. A more successful spline-based technique, discussed by Bliss and Panigirtzoglou (2002), uses a "smoothing" spline. This produces a much better-behaved RND by imposing a penalty function on the choppiness of the spline approximation and not requiring the curve to pass through all of the original points exactly.

Other papers achieve smooth RNDs by positing either a specific returns process (e.g., a jump diffusion) or a specific terminal distribution (e.g., a lognormal) and extracting its

---

[3] The implied volatility literature is voluminous. Poon and Granger (2003) provide an extensive review of this literature, from the perspective of volatility prediction.

parameters from option prices. Nonparametric techniques (e.g., kernel regression) inherently smooth the estimated RND and achieve the same goal.

Several papers in this group, in addition to those described above, are worth mentioning. Bates (1996) used currency options prices to estimate the parameters of a jump-diffusion model for exchange rates, implying out parameters that lead to the best match between the terminal returns distribution under the model and the observed RNDs. Buchen and Kelly (1994) suggested using the principle of Maximum Entropy to establish an RND that places minimal constraints on the data. They evaluated the procedure by simulating options prices and trying to extract the correct density. Bliss and Panigirtzoglou (2002) also used simulated option prices, to compare the performance of smoothing splines versus a mixture of lognormals in extracting the correct RND when prices are perturbed by amounts that would still leave them inside the typical bid-ask spread. They concluded that the spline approach dominates a mixture of lognormals. Bu and Hadri (2007), on the other hand, also used Monte Carlo simulation in comparing the spline technique against a parametric confluent hypergeometric density and preferred the latter.

One might summarize the results from this literature as showing that the implied risk neutral density may be extracted from market option prices using a number of different methods, but none of them is clearly superior to the others. Noisy market option prices and sparse strikes in the available set of traded contracts are a pervasive problem that must be dealt with in any viable procedure. We will select and adapt elements of the approaches used by these researchers to extract the RND from a set of option prices, add a key wrinkle to take account of the bid-ask spread in the market, and then propose a new technique for completing the tails of the distribution.

3. Extracting the Risk Neutral Density from Options Prices, in Theory

In the following, the symbols C, S, X, r, and T all have the standard meanings of option valuation: C = call price; S = time 0 price of the underlying asset; X = exercise price; r = riskless interest rate; T = option expiration date, which is also the time to expiration. P will be the price of a put option. We will also use f(x) = risk neutral probability density function (RND) and F(x) = $\int_{-\infty}^{x} f(z)dz$ = risk neutral distribution function.

The value of a call option is the expected value of its payoff on the expiration date T, discounted back to the present. Under risk neutrality, the expectation is taken with respect to the risk neutral probabilities and discounting is at the risk free interest rate.

(1) $$C = \int_{X}^{\infty} e^{-rT} (S_T - X) f(S_T) dS_T$$

Increasing the exercise price by an amount dX changes the option value for two reasons. First, it narrows the range of stock prices S_T for which the call has a positive payoff. Second, increasing X reduces the payoff by the amount -dX for every S_T at which the option is in the money. The first effect occurs when S_T falls between X and X+dX. The maximum of the lost payoff is just dX, which contributes to option value multiplied by the probability that S_T will end up in that narrow range. So, for discrete dX the impact of the first effect is very small and it becomes infinitesimal relative to the second effect in the limit as dX goes to 0.

These two effects are seen clearly when we take the partial derivative in (1) with respect to X.

$$\frac{\partial C}{\partial X} = \frac{\partial}{\partial X}\left[\int_X^\infty e^{-rT}(S_T - X) f(S_T) dS_T\right]$$

$$= e^{-rT}\left[-(X - X) f(X) + \int_X^\infty -f(S_T) dS_T\right]$$

The first term in brackets corresponds to the effect of changing the range of S_T for which the option is in the money. This is zero in the limit, leaving

$$\frac{\partial C}{\partial X} = -e^{-rT}\int_X^\infty f(S_T) dS_T = -e^{-rT}\left[1 - F(X)\right]$$

Solving for the risk neutral distribution F(X) gives

(2) $$F(X) = e^{rT}\frac{\partial C}{\partial X} + 1$$

In practice, an approximate solution to (2) can be obtained using finite differences of option prices observed at discrete exercise prices in the market. Let there be option prices available for maturity T at N different exercise prices, with X_1 representing the lowest exercise price and X_N being the highest.

We will use three options with sequential strike prices X_{n-1}, X_n, and X_{n+1} in order to obtain an approximation to F(X) centered on X_n.[4]

(3) $$F(X_n) \approx e^{rT}\left[\frac{C_{n+1} - C_{n-1}}{X_{n+1} - X_{n-1}}\right] + 1$$

[4] In general, the differences $(X_n - X_{n-1})$ and $(X_{n+1} - X_n)$ need not be equal, in which case a weighting procedure could be used to approximate $F(X_n)$. In our methodology ΔX is a constant value, because we construct equally spaced artificial option prices to fill in values for strikes in between those traded in the market.

To estimate the probability in the left tail of the risk neutral distribution up to X_2, we approximate $\frac{\partial C}{\partial X}$ at X_2 by $e^{rT} \frac{C_3 - C_1}{X_3 - X_1} + 1$, and the probability in the right tail from X_{N-1} to infinity is approximated by $1 - \left(e^{rT} \frac{C_N - C_{N-2}}{X_N - X_{N-2}} + 1 \right) = -e^{rT} \frac{C_N - C_{N-2}}{X_N - X_{N-2}}$.

Taking the derivative with respect to X in (2) a second time yields the risk neutral density function at X.

(4) $$f(X) = e^{rT} \frac{\partial^2 C}{\partial X^2}$$

The density $f(X_n)$ is approximated as

(5) $$f(X_n) \approx e^{rT} \frac{C_{n+1} - 2C_n + C_{n-1}}{(\Delta X)^2}.$$

Equations (1) - (5) show how the portion of the RND lying between X_2 and X_{N-1} can be extracted from a set of call option prices. A similar derivation can be done to yield a procedure for obtaining the RND from put prices. The equivalent expressions to (2)-(5) for puts are:

(6) $$F(X) = e^{rT} \frac{\partial P}{\partial X}$$

(7) $$F(X_n) \approx e^{rT} \left[\frac{P_{n+1} - P_{n-1}}{X_{n+1} - X_{n-1}} \right]$$

(8) $$f(X) = e^{rT} \frac{\partial^2 P}{\partial X^2}$$

(9) $$f(X_n) \approx e^{rT} \frac{P_{n+1} - 2P_n + P_{n-1}}{(\Delta X)^2}$$

4. Extracting a Risk Neutral Density from Options Market Prices, in Practice

The approach described in the previous section assumes the existence of a set of option prices that are all fully consistent with the theoretical pricing relationship of equation (1). Implementing it with actual market prices for traded options raises several important issues and problems. First, market imperfections in observed option prices must be dealt with carefully or the resulting risk neutral density can have unacceptable features, such as regions in which it is negative. Second, some way must be found to complete the tails of the RND beyond the range from X_2 to X_{N-1}. This section will review several approaches that have been used in the literature to obtain the middle portion of the RND from market option prices, and will describe the technique we adopt here. The next section will add the tails.

We will be estimating RNDs from the daily closing bid and ask prices for Standard and Poor's 500 Index options. S&P 500 options are particularly good for this exercise because the underlying index is widely accepted as the proxy for the U.S. "market portfolio," the options are very actively traded on the Chicago Board Options Exchange, and they are cash-settled with European exercise style. S&P 500 options have major expirations quarterly, on the third Friday of the months of March, June, September and December. This will allow us to construct time series of RNDs applying to the value of the S&P index on each expiration date. The data set will be described in further detail below. Here we will take a single date, January 5, 2005, selected at random, to illustrate extraction of an RND in practice.

Interpolation and Smoothing

The available options prices for 1/5/2005 are shown in Table 1. The index closed at 1183.74 on that date, and the March options contracts expired 72 days later, on March 18, 2005. Strike prices ranged from 1050 to 1500 for calls, and from 500 to 1350 for puts. Bid-ask spreads were relatively wide: 2 points for contracts trading above 20 dollars down to a minimum of 0.50 in most cases even for the cheapest options. This amounted to spreads of more than 100% of the average price for many deep out of the money contracts. It is customary to use either transactions prices or the midpoints of the quoted bid-ask spreads as the market's option prices. Options transactions occur irregularly in time and only a handful of strikes have frequent trading, even for an actively traded contract like S&P 500 index options. Use of transactions data also requires obtaining synchronous prices for the underlying. By contrast, bids and offers are quoted continuously for all traded strikes, whether or not trades are occurring. We will begin by taking the average of bid and ask as the best available measure of the option price. We then modify the procedure to make use of the full spread in the smoothing and interpolation stage.

Equations (3) and (7) show how to estimate the probability distribution using a centered difference to compute the slope and the distribution at X_n. In Figure 1, we have used uncentered differences, $C_n - C_{n-1}$ and $P_n - P_{n-1}$ simply for illustration, to construct probability distributions from the average call and put price quotes shown in Table 1. The distribution from the puts extends further to the left and the one from the calls extends further to the right, but in the middle range where they overlap, the values are quite close together. There are some discrepancies, notably around 1250, where the cumulative call probability is 0.698 and the put probability is 0.776, but the more serious problem is around 1225, where the fitted probability distribution from call prices is non-monotonic.

Figure 2 plots the risk neutral densities corresponding to the distribution functions displayed in Figure 1. These are clearly unacceptable as plausible estimates of the true density function. Both RNDs have ranges of negative values, and the extreme fluctuations in the middle portion and sharp differences between call and put RNDs violate our prior beliefs that the RND should be fairly smooth and the same expectations should govern pricing of both calls and puts.

Looking at the prices in Table 1, it is clear that there will be problems with out of the money puts. Except at 800, there is no bid for puts at any strike below 925 and the ask price is unchanged over multiple contiguous strikes, making the average price equal for different exercise prices. From (9), the estimated RND over these regions will be 0, implying no possibility that the S&P could end up there at expiration. A similar situation occurs for out of the money calls between 1400 and 1500. Moreover, the single 0.10 bid for puts at $X = 800$ produces an average put price higher than that for the next higher strike, which violates the static no-arbitrage condition that a put with a higher strike must be worth more than one with a lower strike. This leads to a region in which the implied risk neutral density is negative. However, it is not obvious from the prices in Table 1 what the problem is that produces the extreme choppiness and negative densities around the at the money index levels between 1150 and 1250.

Table 1 and Figure 1 show that even for this very actively traded index option, the available strikes are limited and the resulting risk neutral distribution is a coarse step function. The problem would be distinctly worse for individual traded stock options whose available strikes are considerably less dense than this. This suggests the use of an interpolation technique to fill in intermediate values between the traded strike prices and to smooth out the risk neutral distribution.

Cubic spline interpolation is a very common first choice as an interpolation tool. Figure 3 shows the spline-interpolated intermediate option prices for our calls and puts. To the naked eye, the curves look extremely good, without obvious bumps or wiggles between the market prices, indicated by the markers. Yet these option prices produce the RNDs shown in Figure 4, with erratic fluctuations around the at the money stock prices, large discrepancies between RNDs from calls and puts, and negative portions in both curves. The problem is that cubic spline interpolation generates a curve that is forced to go through every observed price, which has the effect of incorporating all of the noise due to market microstructure and other imperfections into the RND.

David Shimko (1993) proposed transforming the market option prices into implied volatility (IV) space before interpolating, then retransforming the interpolated curve back to price space to compute a risk neutral distribution. This procedure does not assume that the Black-Scholes model holds for these option prices. It simply uses the Black-Scholes equation as a computational device to transform the data into a space which is more conducive to the kind of smoothing one wishes to do.

Consider the transformed option values represented by their BS IVs. Canonical Black-Scholes would require all of the options to have the same IV. If this constraint were imposed and the fitted IVs transformed back into prices, by construction the resulting risk neutral density would be lognormal, and hence well-behaved. But because of the well-known volatility smile, or skew in this market, the new prices would be systematically different from the observed market prices, especially in the left tail. Most option traders do not use the canonical form of the BS model, but instead use "practitioner Black-Scholes," in which each option is allowed to have its own distinct implied volatility. Despite the theoretical inconsistency this introduces, the empirical volatility smile/skew is quite smooth and not too badly sloped, so it works well enough.

Considerable research effort has been devoted to finding arbitrage-free theoretical models based on nonlognormal returns distributions, that produce volatility smiles resembling those found empirically. Inverting those (theoretical) smiles will also lead to option prices that produce well-behaved RNDs. Of course, if market prices do not obey the alternative theoretical model due to market noise, transforming through implied volatility space won't cure the problem.

To moderate the effects of market imperfections in option prices, a smooth curve is fitted to the volatility smile/skew by least squares. Shimko used a simple quadratic function, but we prefer to allow greater flexibility with a higher order polynomial.

Applying a cubic spline to interpolate the volatility smile still produces bad results for the fitted RND. The main reason for this is that an n-th degree spline constructs an interpolating curve consisting of segments of n-th order polynomials joined together at a set of "knot" points. At each of those points, the two curve segments entering from the left and the right are constrained to have the same value and the same derivatives up to order n-1. Thus a cubic spline has no discontinuities in the level, slope or 2nd derivative, meaning there will be no breaks, kinks, or even visible changes in curvature at its knot points. But when the interpolated IV curve is translated back into option strike-price space and the RND is constructed by taking the second derivative as in (5), the discontinuous 3rd derivative of the IV curve becomes a discontinuous first derivative--a kink--in the RND. The simple solution is just to interpolate with a 4th order spline or higher.[5]

[5] As mentioned above, some researchers plot the IV smile against the option deltas rather than against the strike prices, which solves this problem automatically. Applying a cubic spline in delta-IV space produces a curve that is smooth up to 2nd order in terms of the partial derivative of option price, which makes it smooth up to 3rd order in the price itself, eliminating any kinks in the RND.

The other problem with using a standard n-th degree spline as an interpolating function is that it must pass through every knot point, which forces the curve to incorporate all pricing noise into the RND. Since with K knot points, there will be K+n+1 parameters to fit, this also requires applying enough constraints to the curve at its endpoints to allow all of the parameters to be identified with only K data points.

Previous researchers have used a "smoothing spline" that allows a tradeoff between how close the curve is to the observed data points--it no longer goes through them exactly--and how well its shape conforms to the standard spline constraint that the derivatives of the spline curve should be smooth across the knot points. For any given problem, the researcher must choose how this tradeoff is resolved by setting the value of a smoothness parameter.[6]

We depart somewhat from previous practice in this area. We have found that fitted RNDs behave very well using interpolation with just a 4th order polynomial--essentially a 4th degree spline with no knots. Additional degrees of freedom, that allow the estimated densities to take more complex shapes, can be added either by fitting higher order polynomials or by adding knots to a 4th order spline. In this exercise, we found very little difference from either of these modifications. We therefore have done all of the interpolation for our density estimation using 4th order splines with a single knot point placed at-the-money.

Looking again at Table 1, we see that many of the bid and ask quotes are for options that are either very deep in the money or very deep out of the money. For the former case, the effect of optionality is quite limited, such that the IV might range from 12.9% to 14.0% within the bid-ask spread. For the lowest strike call, there is no IV at the bid price, because it is below the no-arbitrage minimum call price. The IV at the ask is 15.6%, while the IV at the midpoint, which is what goes into the calculations, is 11.8%. In addition to the wide bid-ask spreads, there is little or no trading in deep in the money contracts. On this day, no 1050 or 1075 strike calls were traded at all, and only 3 1150 strike calls changed hands. Most of the trading is in at the money or out of the money contracts.

But out of the money contracts present their own data problems, because of extremely wide bid-ask spreads relative to their prices. The 925 strike put, for example, would have an IV of 22.3% at its bid price of 0.20 and 26.2% at the ask price of 0.70. Setting the IV for this option at 24.8% based on the mid-price of 0.45 is clearly rather arbitrary. One reason the spread is so wide is that there is very little trading of deep out of the money contracts. On this date, the only trades in puts with strikes of 925 or below were 5 contracts at a strike of 850, for a total option premium of no more than a couple hundred dollars. It is obvious that the quality of information about the risk neutral density that can

[6] The procedure imposes a penalty function on the integral of the second derivative of the spline curve to make the fitted curve smoother. The standard smoothing spline technique still uses a knot at every data point, so it requires constraints to be imposed at the endpoints. See Bliss and Panigirtzoglou (2002), Appendix A, for further information about this approach.

be extracted from the posted quotes on options that don't trade in the market may be quite limited.

These observations suggest that it is desirable to limit the range of option strikes that are brought into the estimation process, eliminating those that are too deep in or out of the money. Also, since most trading is in at the money and somewhat out of the money contracts, we can broaden the range with usable data if we combine calls and puts together. The CBOE does this in their calculation of the VIX index, for example, combining calls and puts but using only out of the money contracts.

To incorporate these ideas into our methodology, we first discard all options whose bid prices are less than 0.50. On this date, this eliminates calls with strikes of 1325 and above, and puts with strikes of 925 and below. Next we want to combine calls and puts, using the out of the money contracts for each. But from Table 1, with the current index level at 1183.74, if we simply use puts with strikes up to 1180 and calls with strikes from 1190 to 1300, there will be a jump from the put IV of 14.2% to the call IV of 12.6% at the break point. To smooth out the effect of this jump at the transition point, we blend the call and put IVs in the region around the at the money index level. We have chosen a range of 20 points on either side of the current index value S_0 in which the IV will be set to a weighted average of the IVs from the calls and the puts.[7]

Let X_{low} be the lowest traded strike such that $(S_0 - 20) \leq X_{low}$ and X_{high} be the highest traded strike such that $X_{high} \leq (S_0 + 20)$. For traded strikes between X_{low} and X_{high} we use a blended value between $IV_{put}(X)$ and $IV_{call}(X)$, computed as

(10) $$IV_{blend}(X) = w\, IV_{put}(X) + (1-w)\, IV_{call}(X)$$

where $$w = \frac{X_{high} - X}{X_{high} - X_{low}}$$

In this case, we take put IVs for strikes up to1150, blended IVs for strikes 1170 to 1200, and call IVs for strikes from 1205 up. Figure 5 plots the raw IVs from the traded options with markers and the interpolated IV curve computed from calls and puts whose bid prices are at least 0.50, as just described.

[7] The choice of a 40 point range over which to blend the put and call IVs is arbitrary, but we believe that the specific choice has little impact on the overall performance of the methodology. On Jan. 5, 2005, the discrepancy between the two IVs is about .015 in this range, which becomes distributed over the 40 point range of strikes at the rate of about 0.0004 per point. The effect on the fitted RND will be almost entirely concentrated around the midpoint, and it will be considerably smoother than if no adjustment were made and the IV simply jumped from the put value to the call value for the at the money strike. A reasonable criterion in setting the range for IV blending would be to limit it to the area before the IVs from the two sets begin to diverge, as Figure 5 illustrates happens when one of them gets far enough out of the money.

This procedure produces an implied risk neutral density with a very satisfying shape, based on prior expectations that the RND should be smooth. Even so, there might be some concern that we have smoothed out too much. We have no reason to rule out minor bumps in the RND, that could arise when an important dichotomous future event is anticipated, such as the possibility of a cut in the Federal Reserve's target interest rate, or alternatively, if there are distinct groups in the investor population with sharply divergent expectations.

I have explored increasing flexibility by fitting 4th order splines using 3 knots with one at the midpoint and the others 20 points above and below that price. The choice of how many knots to use and where to place them allows considerable latitude for the user. But we will see shortly that at least in the present case, it makes very little difference to the results.

Incorporating Market Bid-Ask Spreads

The spline is fitted to the IV observations from the market by least squares. This applies equal weights to the squared deviation between the spline curve and the market IV evaluated at the midpoint of the bid-ask spread at all data points, regardless of whether the spline would fall inside or outside the quoted spread. Given the width of the spreads, it would make sense to be more concerned about cases where the spline fell outside the quoted spread than those remained within it.

To take account of the bid-ask spread, we apply a weighting function to increase the weighting of deviations falling outside the quoted spread relative to those that remain within it. We adapt the cumulative normal distribution function to construct a weighting function that allows weights between 0 to 1 as a function of a single parameter σ.

(11)
$$w(IV) = \begin{cases} N[\,IV - IV_{Ask}\,,\,\sigma] & \text{if } IV_{Midpoint} \le IV \\ N[\,IV_{Bid} - IV\,,\,\sigma] & \text{if } IV \le IV_{Midpoint} \end{cases}$$

Figure 6 plots an example of this weighting function for three values of σ. Implied volatility is on the x axis, with the vertical solid lines indicating a given option's IV values at the market's bid, ask, and midprice, 0.1249, 0.1331, and 0.1290, respectively. These values are obtained by applying the spline interpolation described above separately to the three sets of IVs, from the bid prices, the ask prices and the midprices at each traded strike level. In the middle range where call and put IVs are blended according to equation (10) the bid and ask IV curves from calls and puts are blended in the same way before the interpolation step.

Settting σ to a very high value like 100 assigns (almost) equal weights of 0.5 to all squared deviations between the IV at the midpoint and the fitted spline curve at every strike. This is the standard approach that does not take account of the bid-ask spread. With $\sigma = 0.005$, all deviations are penalized, but those falling well outside the quoted spread are weighted about 3 times more heavily than those close to the midprice IV. Setting $\sigma = 0.001$ puts very little weight on deviations that are within the spread and close to the midprice IV, while assigning full weight to nearly all deviations falling outside the spread. This is our preferred weighting pattern to make use of the information contained in the quoted spread in the market.

Figure 7 illustrates the effect of changing the degree of the polynomial, the number of knot points and the bid-ask weighting parameter used in the interpolation step. Lines in gray show densities constructed by fitting polynomials of degree 4, 6, and 8, with no knots and equal weighting of all squared deviations. The basic shape of the three curves is close, but higher order polynomials allow greater flexibility in the RND. This allows it to fit more complex densities, but also increases the impact of market noise. Consider the left end of the density. The missing portion of the left tail must be attached below 950 but it is far from clear how it should look to match the density obtained either from the 8th degree polynomial, which slopes sharply downward at that level, or from the 6th degree polynomial, which has a more reasonable slope at that point, but the estimated density is negative.

By contrast, the 4th order polynomial and all three of the spline functions produce very reasonably shaped RNDs that are so close together that they can not be distinguished in the graph. Although these plots are for a single date, I have found similar results on nearly every date for which this comparison was done, which supports the choice of a 4th order spline with a single knot and with a very small relative weight on deviations that fall within the bid-ask spread in order to extract risk neutral densities from S&P 500 index options.

Summary

The following steps summarize our procedure for extracting a well-behaved risk neutral density from market prices for S&P 500 index options, over the range spanned by the available option strike prices.

1. Begin with bid and ask quotes for calls and puts with a given expiration date.

2. Discard quotes for very deep out of the money options. We required a minimum bid price of $0.50 for this study.

3. Combine calls and puts to use only the out of the money and at the money contracts, which are the most liquid.

4. Convert the option bid, ask and midprices into implied volatilities using the Black-Scholes equation. To create a smooth transition from put to call IVs, take weighted

averages of the bid, ask and midprice IVs from puts and calls in a region around the current at the money level, using equation (10).

5. Fit a spline function of at least 4th order to the midprice implied volatilities by minimizing the weighted sum of squared differences between the spline curve and the midprice IVs. The weighting function shown in equation (11) downweights deviations that lie within the market's quoted bid-ask spread relative to those falling outside it. The number of knots should be kept small, and their optimal placement may depend on the particular data set under consideration. In this study we used a 4th order spline with a single knot at the money.

6. Compute a dense set of interpolated IVs from the fitted spline curve and then convert them back into option prices.

7. Apply the procedure described in Section 3 to the resulting set of option prices in order to approximate the middle portion of the RND.

These steps produce an empirical RND over the range between the lowest and highest strike price with usable data. The final step is to extend the density into the tails.

5. Adding Tails to the Risk Neutral Density

The range of strike prices $\{X_1, X_2, ..., X_N\}$ for which usable option prices are available from the market or can be constructed by interpolation does not extend very far into the tails of the distribution. The problem is further complicated by the fact that what we are trying to approximate is the market's aggregation of the individual risk neutralized subjective probability beliefs in the investor population. The resulting density function need not obey any particular probability law, nor is it even a transformation of the true (but unobservable) distribution of realized returns on the underlying asset.

We propose to extend the empirical RND by grafting onto it tails drawn from a suitable parametric probability distribution in such a way as to match the shape of the estimated RND over the portion of the tail region for which it is available. The first question is which parametric probability distribution to use. Some of the earlier approaches to this problem implicitly assume a distribution. For example, the Black-Scholes implied volatility function can be extended by setting $IV(X) = IV(X_1)$ for all $X < X_1$ and $IV(X) = IV(X_N)$ for all $X > X_N$, where $IV(\cdot)$ is the implied volatility from the Black-Scholes model.[8] This forces the tails to be lognormal. Bliss and Panigirtzoglou (2004) do something similar by employing a smoothing spline for the middle portion of the distribution but constraining it to become linear outside the range of the available strikes. Given the extensive empirical evidence of fat tails in returns distributions, constraining

[8] See, for example Jiang and Tian (2005).

the tails of the RND to be lognormal is unlikely to be satisfactory in practice if one is concerned about modeling tail events accurately.

Fortunately, similar to the way the Central Limit Theorem makes the normal a natural choice for modeling the distribution of the sample average from an unknown distribution, the Extreme Value distribution is a natural candidate for the purpose of modeling the tails of an unknown distribution. The Fisher-Tippett Theorem proves that under weak regularity conditions the largest value in a sample drawn from an unknown distribution will converge in distribution to one of three types of probability laws, all of which belong to the generalized extreme value (GEV) family.[9] We will therefore use the GEV distribution to construct tails for the RND.

The standard Generalized Extreme Value distribution has one parameter ξ, which determines the tail shape.

GEV Distribution function:

(12) $$F(z) = \exp[-(1 + \xi z)^{-1/\xi}]$$

The value of ξ determines whether the tail comes from the Fréchet distribution with fat tails relative to the normal ($\xi > 0$), the Gumbel distribution with tails like the normal ($\xi = 0$), or the Weibull distribution ($\xi < 0$) with finite tails that to do not extend out to infinity.

Two other parameters, μ and σ, can be introduced to set location and scale of the distribution, by defining

(13) $$z = \frac{S_T - \mu}{\sigma}$$

Thus we have three GEV parameters to set, which allows us to impose three conditions on the tail. We will use the expressions $F_{EVL}(\cdot)$ and $F_{EVR}(\cdot)$ to denote the approximating GEV distributions for the left and right tails, respectively, with $f_{EVL}(\cdot)$ and $f_{EVR}(\cdot)$ as the corresponding density functions, and the same notation without the L and R subscripts when referring to both tails without distinction. $F_{EMP}(\cdot)$ and $f_{EMP}(\cdot)$ will denote the estimated empirical risk neutral distribution and density functions.

Let $X(\alpha)$ denote the exercise price corresponding to the α-quantile of the risk neutral distribution. That is, $F_{EMP}(X(\alpha)) = \alpha$. We first choose the value of α where the GEV tail is to begin, and then a second, more extreme point on the tail, that will be used in

[9] Specifically, let $x_1, x_2, \ldots$ be an i.i.d. sequence of draws from some distribution F and let M_n denote the maximum of the first n observations. If we can find sequences of real numbers a_n and b_n such that the sequence of normalized maxima $(M_n - b_n)/a_n$ converges in distribution to some non-degenerate distribution H(x), i.e., $P((M_n - b_n)/a_n \leq x) \rightarrow H(x)$ as $n \rightarrow \infty$ then H is a GEV distribution. The class of distribution functions that satisfy this condition is very broad, including all of those commonly used in finance. See Embrechts, et al (1997) or McNeil, et al (2005) for further detail.

matching the GEV tail shape to that of the empirical RND. These values will be denoted α_{0R} and α_{1R}, respectively, for the right tail and α_{0L} and α_{1L} for the left.

The choice of α_0 and α_1 values is flexible, subject to the constraint that we must be able to compute the empirical RND at both points, which requires $X_2 \leq X(\alpha_{1L})$ and $X(\alpha_{1R}) \leq X_{N-1}$. However, the GEV will fit the more extreme tail of an arbitrary distribution better than the near tail, so there is a tradeoff between data availability and quality, which would favor less extreme values for α_0 and α_1, versus tail fit, which would favor more extreme values.

Consider first fitting a GEV upper tail for the RND. The first condition to be imposed is that the total probability in the tail must be the same for the RND and the GEV approximation. We also want the GEV density to have the same shape as the RND in the area of the tail where the two overlap, so we use the other two degrees of freedom to set the two densities equal at α_{0R} and α_{1R}.

The three conditions for the right tail are shown in equations (14a-c):

(14a) $$F_{EVR}(X(\alpha_{0R})) = \alpha_{0R}$$

(14b) $$f_{EVR}(X(\alpha_{0R})) = f_{EMP}(X(\alpha_{0R}))$$

(14c) $$f_{EVR}(X(\alpha_{1R})) = f_{EMP}(X(\alpha_{1R}))$$

The GEV parameter values that will cause these conditions to be satisfied can be found easily using standard optimization procedures.

Fitting the left tail of the RND is slightly more complicated than the right tail. Since the GEV is the distribution of the maximum in a sample, its left tail relates to probabilities of small values of the maximum, rather than to extreme values of the sample minimum, i.e., the left tail. To adapt the GEV to fitting the left tail, we must reverse it left to right, by defining it on -z. That is, z values must be computed from (15) in place of (13):

(15) $$z = \frac{(-\mu_L) - S_T}{\sigma}$$

where μ_L is the (positive) value of the location parameter for the left tail GEV. (The optimization algorithm will return the location parameter $\mu \equiv -\mu_L$ as a negative number.)[10]

---

[10] The procedure as described works well for fitting tails to an RND that is defined on positive X values only, as it is when X refers to an asset price S_T, or a simple gross return S_T/S_0. Fitting an RND in terms of

The optimization conditions for the left tail become

(16a) $$F_{EVL}(-X(\alpha_{0L})) = 1 - \alpha_{0L}$$

(16b) $$f_{EVR}(-X(\alpha_{0L})) = f_{EMP}(X(\alpha_{0L}))$$

(16c) $$f_{EVR}(-X(\alpha_{1L})) = f_{EMP}(X(\alpha_{1L}))$$

Our initial preference was to connect the left and right tails at α_0 values of 5% and 95%, respectively. However, for the S&P 500 index options in the sample that will be analyzed below, market prices for options with the relevant exercise prices were not always available for the left tail and rarely were for the right tail. We have therefore chosen default values of $\alpha_{0L} = 0.05$ and $\alpha_{0R} = 0.92$, with $\alpha_{1L} = 0.02$ and $\alpha_{1R} = 0.95$ as the more remote connection points. In cases where data were not available for these α values, we set $\alpha_{1L} = F_{EMP}(X_2)$, the lowest connection point available from the data, and $\alpha_{0L} = \alpha_{1L} + 0.03$. For the right tail, $\alpha_{1R} = F_{EMP}(X_{N-1})$, and $\alpha_{0R} = \alpha_{1R} - 0.03$.

On Jan. 5, 2005, the 5% and 2% quantiles of the empirical RND fell at 1044.00 and 985.50, respectively, and the 95% and 92% right-tail quantiles were 1271.50 and 1283.50, respectively.[11] The fitted GEV parameters that satisfied equations (14) and (16) were as follows:

Left tail: $\mu = 1274.60$ $\sigma = 91.03$ $\xi = -0.112$

Right tail: $\mu = 1195.04$ $\sigma = 36.18$ $\xi = -0.139$

Figure 8 plots three curves: the middle portion of the empirical RND extracted from the truncated set of options prices with interpolation using a 4th degree spline with one knot at the money and bid-ask weighting parameter $\sigma = .001$, as shown in Figure 7, and the two GEV distributions whose tails have been matched to the RND at the four connection points. As the Figure illustrates, the GEV tail matches the empirical RND very closely in

---

log returns, however, raises a problem that it may not be possible to fit a good approximating GEV function on the same support as the empirical RND. This difficulty can be dealt with by simply adding a large positive constant to every X value to shift the empirical RND to the right for fitting the tails, and then subtracting it out afterwards, to move the completed RND back to the right spot on the x-axis.

[11] With finite stock price increments in the interpolation, these quantiles will not fall exactly on any X_n . We therefore choose n at the left tail connection points such that $X_{n-1} \leq X(\alpha) < X_n$ and set the actual quantiles α_{0L} and α_{1L} equal to the appropriate actual values of the empirical risk neutral distribution and density at X_n. Similarly, the right connection points are set such that $X_{n-1} < X(\alpha) \leq X_n$.

the region of the 5% and 92% tails. Figure 9 shows the resulting completed RND with GEV tails.

6. Estimating the Risk Neutral Density for the S&P 500 from S&P 500 Index Options

We applied the methodology described above to fit risk neutral densities for the Standard and Poor's 500 stock index using S&P 500 index call and put options over the period January 4, 1996 - February 20, 2008. In this section we will present interesting preliminary results on some important issues, obtained from analyzing these densities. The purpose is to illustrate the potential of this approach to generate valuable insights about how investors' information and risk preferences are incorporated in market prices. The issues we consider are complex and we will not attempt to provide in-depth analysis of them in this paper. Rather, we offer a small set of what we hope are tantalizing "broad brush" results that suggest directions in which further research along these lines is warranted. Specifically, we first examine the moments of the fitted RNDs and compare them to the lognormal densities assumed in the Black-Scholes model. We then look at how the RND behaves dynamically, as the level of the underlying index changes.

Data Sample

Closing bid and ask option prices data were obtained from Optionmetrics through the WRDS system. The RND for a given expiration date is extracted from the set of traded options with that maturity, and each day's option prices provide an updated RND estimate for the same expiration date. We focus on the quarterly maturities with expirations in March, June, September, and December, which are the months with the most active trading interest.[12] The data sample includes option prices for 49 contract maturities and 2761 trading days.

We construct RNDs, updated daily, for each quarterly expiration, beginning immediately after the previous contract expires and ending when the contract has less than 2 weeks remaining to maturity. Very short maturity contracts were eliminated because we found that their RNDs are often badly behaved. This may be partly due to price effects from trading strategies related to contract expiration and rollover of hedge positions into later expirations. Also, the range of strikes for which there is active trading interest in the market gets much narrower as expiration approaches.

We computed Black-Scholes IVs using the closing bid and ask prices reported by Optionmetrics. Optionmetrics was also the source for the riskless rate and dividend yield

---

[12] The CBOE lists contracts with maturities in the next 3 calendar months plus 3 more distant months from the March-June-September-December cycle, meaning that off-month contracts such as April and May are only introduced when the time to maturity is less than 3 months.

data, which are also needed in calculating forward values for the index on the option maturity dates.[13]

Table 2 provides summary information on the data sample and the estimated tail parameters. During this period, the S&P index ranged from a low of just under 600 to a high of 1565.20, averaging 1140.60. Contract maturities were between slightly over 3 months down to 14 days, with an average value of about 54 days.

The number of market option prices available varied from day to day and some of those for which prices were reported were excluded because implied volatilities could not be computed (typically because the option price violated a no-arbitrage bound). The numbers of usable calls and puts averaged about 46 and 42 each day, respectively. We eliminated those with bid prices in the market less than \$0.50. The excluded deep out of the money contracts are quite illiquid and, as Table 1 shows, their bid-ask spreads are very wide relative to the option price. On average about 38 calls and 33 puts were used to fit a given day's RND, with a minimum of 6 puts and 7 calls. Their implied volatilities averaged around 25%, but covered a very wide range of values.

The tail parameters reported in the table relate to risk neutral densities estimated on gross returns, defined as S_T / S_0, where S_0 is the current index level and S_T is the index on the contract's expiration date. This rescaling makes it possible to combine RNDs from different expirations so that their tail properties can be compared. Under Black-Scholes assumptions, these simple returns should have a lognormal distribution.

For the left tail, if sufficient option price data are available, the connection point is set at the index level where the empirical RND has cumulative probability of $\alpha_{0L} = 5\%$. This averaged 0.8672, i.e., a put option with that strike was about 13% out of the money. The mean value of the fitted left tail shape parameter ξ was 0.0471, which makes the left tail shape close to the normal on average, but with a fairly large standard deviation. Note that this does not mean the RND is not fat-tailed relative to the normal, as we will see when we look at its excess kurtosis in Table 3, only that the extreme left tail of the RND defined on simple returns is not fat on average. Indeed, since the RND defined on gross returns is bounded below by 0, the true left tail must be thin-tailed relative to the normal, asymptotically

The right connection point averaged 1.0900, i.e., where a call was 9% out of the money. The tail shape parameter ξ was negative for the right tail, implying a short-tailed distribution with a density that hits zero at a finite value. This result was very strong: although the fitted values for ξ varied over a fairly wide range, with a standard deviation of 0.0707, only 1 out of 2761 ξ estimates for the right tail was positive. Comparing the σ estimates for the left and right tails, we see that the typical GEV approximations

[13] Optionmetrics interpolates U.S. dollar LIBOR to match option maturity and converts it into a continuously compounded rate. The projected dividends on the index are also converted to a continuous annual rate. See the Optionmetrics Manual (2003) for detailed explanations of how Optionmetrics handles the data.

generally resemble those shown for January 5, 2005 in Figures 8 and 9, with the left tail coming from a substantially wider distribution than the right tail.

Moments of the Risk Neutral Density

Table 3 displays summary statistics on the moments of the fitted S&P 500 risk neutral densities. The table, showing the mean, standard deviation, and several quantiles of the distribution of the first four moments of the fitted densities within the set of 2761 estimated RNDs, provides a number of interesting results.

The mean risk neutralized expected return including dividends was 0.61%, over time horizons varying from 3 months down to 2 weeks. At annualized rates, this was 4.05%, but with a standard deviation of 1.89%. The quantile results indicate that the range of expected returns was fairly wide. Perhaps more important is how the expected return option traders expected to earn compared to the riskless rate. Under risk neutrality, the expected return on any security, including the stock market portfolio, should be equal to the riskless interest rate, but the third row of Table 3 shows that on average, option traders expected a risk neutralized return 21 basis point below the riskless rate (using LIBOR as the proxy for that rate). The discrepancy was distributed over a fairly narrow range, however, with more than 95% of the values between -1% and +1%.

Skewness of the RND defined over returns was strongly negative. In fact, the skewness of the RND was negative on every single day in the sample. Under Black-Scholes assumptions, the distribution of gross returns is lognormal and risk neutralization simply shifts the density to the left so that its mean becomes the riskless rate. The skewness result in Table 3 strongly rejects the hypothesis that the risk neutral density is consistent with the standard model. Kurtosis was well over 3.0, indicating the RNDs were fat-tailed relative to the normal, although the nonzero skewness makes this result difficult to interpret clearly.

To explore these results a little further, we converted the RNDs defined on terminal index levels to RNDs for log returns, defined as $r = \log(S_T/S_0)$.[14] This would yield a normal distribution if returns were lognormal. The results for skewness and excess kurtosis are shown in Table 3 for comparison, and they confirm what we have seen for gross returns. The RND defined on log returns is even more strongly left-skewed and excess kurtosis is

[14] Let x be a continuous r.v. with density $f_x(\,.\,)$. Let $y = g(x)$ be a one-to-one transformation of x such that the derivative of $x = g^{-1}(y)$ with respect to y is continuous. Then $Y = g(X)$ is a continuous r.v. with density

$$f_Y(y) \quad = \quad \left|\frac{d}{dy} g^{-1}(y)\right| f_X(g^{-1}(y))$$

In our case, $r = g(S) = \ln(S/S_0)$. Therefore, $RND_r(r) = S \times RND_S(S)$, where $r = \ln(S/S_0)$.

increased. The RND was fat-tailed relative to the normal on every single day in the sample.

The Dynamic Behavior of the S&P 500 Risk Neutral Density

How does the RND behave when the underlying Standard and Poor's 500 index moves? The annualized excess return shown in Table 3 implies that the mean of the RND (in price space) is approximately equal to the forward price of the index on average. A reasonable null hypothesis would therefore be that if the forward index value changes by some amount ΔF, the whole RND will shift left or right by that amount at every point. This does not appear to be true.

Table 4 reports the results of regressing the changes in quantiles of the RND on the change in the forward value of the index. Regression (17) was run for each of 11 quantiles, Q_j, j=1,...11, of the risk neutral densities.

(17) $$Q_j(t) = a + b\,\Delta F(t)$$

Under the null hypothesis that the whole density shifts up or down by $\Delta F(t)$ as the index changes, the coefficient b should be 1.0 for all quantiles.

When (17) is estimated on all observations, all b coefficients are positive and highly significant, but they show a clear negative and almost perfectly monotonic relationship between the quantile and the size of b. When the index falls, the left end of the curve drops by more than the change in the forward index and the right end moves down by substantially less. For example, a 10 point drop in the forward index leads to about a 14 point drop in the 1% and 2% quantiles, but the upper quantiles, 0.90 and above, go down less than 8 points. Similarly, when the index rises the lower quantiles go up further than the upper quantiles. Visually, the RND stretches out to the left when the S&P drops, and when the S&P rises the RND tends to stack up against its relatively inflexible upper end.

The next two sets of results compare the behavior of the quantiles between positive and negative returns. Although the same difference in the response of the left and right tails is present in both cases, it is more pronounced when the market goes down than when it goes up. To explore whether a big move has a different impact, the last two sets of results in Table 4 report regression coefficients fitted only on days with large negative returns, below -1.0%, or large positive returns greater than +1.0%. When the market falls sharply, the effect on the left tail is about the same as the overall average response to both up and down moves, but the extreme right tail moves distinctly less than for a normal day. By contrast, if the market rises more than 1.0%, the left tail effect is attenuated while the right tail seems to move somewhat more than for a normal day.

These interesting and provocative results on how the RND responds to and reflects the market's changing expectations and (possibly) risk attitudes as prices fluctuate in the market warrant further investigation. One potentially important factor here is that the biggest differences are found in the far tails of the RND, in the regions where the empirical RND has been extended with GEV tails. What we are seeing may be a result of changes in the shape of the empirical RND at its ends when the market makes a big move, which the GEV tails then try to match. Note, however, that the empirically observed portion of the RND for the full sample shows the strong monotonic coefficient estimates throughout its full range, so the patterns revealed in Table 4 are clearly more than simply artifacts of the tail fitting procedure.

7. Concluding Comments

We have proposed a comprehensive approach for extracting a well-behaved estimate of the risk neutral density over the price or return of an underlying asset, using the market prices of its traded options. This involves two significant technical problems: first, how best to obtain a sufficient number of valid option prices to work with, by smoothing the market quotes to reduce the effect of market noise, and interpolating across the relatively sparse set of traded strike prices; and second, how to complete the density functions by extending them into the tails. We explored several potential solutions for the first problem and settled on converting market option price quotes into implied volatilities, smoothing and interpolating them in strike price-implied volatility space, converting back to a dense set of prices, and then applying the standard methodology to extract the middle portion of the risk neutral density. We then addressed the second problem by appending left and right tails from a Generalized Extreme Value distribution in such a way that each tail contains the correct total probability and has a shape that approximates the shape of the empirical RND in the portion of the tail that was available from the market.

Although the main concentration in this paper has been on developing the estimation technology, the purpose of the exercise is ultimately to use the fitted RND functions to learn more about how the market prices options, how it responds to the arrival of new information, and how market risk preferences behave and vary over time. We presented results showing that the risk neutral density for the S&P 500 index, as reflected in its options, is far from the lognormal density assumed by the Black-Scholes model--it is strongly negatively skewed and fat-tailed relative to the (log)normal. We also found that when the underlying index moves, the RND not only moves along with the index, but it also changes shape in a regular way, with the left tail responding much more strongly than the right tail to the change in the index.

These results warrant further investigation, for the S&P 500 and for other underlying assets that have active options markets. The following is a selection of such projects that are currently under way.

The Federal Reserve announces its Federal funds interest rate target and policy decisions at approximately 2:15 in the afternoon at the end of its regular meeting, about every 6 weeks. This is a major piece of new information and the market's response is immediate and often quite extreme. Using intraday options data, it is possible to estimate real time RNDs that give a very detailed picture of how the market's expectations and risk preferences are affected by the information release.

The volatility of the underlying asset is a very important input into all modern option pricing models, but volatility is hard to predict accurately and there are a number of alternative techniques in common use. There are also a number of index-based securities that are closely related to one another and should therefore have closely related volatilities. The RND provides insight into what the market's expected volatility is, and how it is connected to other volatility measures, like realized historical volatility, volatility estimated from a volatility model such as GARCH, realized future volatility over the life of the option, implied volatility from individual options or from the VIX index, volatility of S&P index futures prices, implied volatility from futures options, volatility of the SPDR tracking ETF, etc.

Yet another important issue involves causality and predictive ability of the risk neutral density. Does the information contained in the RND predict the direction and volatility of future price movements, or does it lag behind and follow the S&P index or the S&P futures price?

We hope and anticipate that the procedure we have developed here can be put to work in these and other projects, and will ultimately generate valuable new insights into the behavior of financial markets.

References

Ait-Sahalia Y. and A.W. Lo. (1998). "Nonparametric Estimation of State-Price Densities Implicit in Financial Asset Prices." Journal of Finance 53, 499-547.

Ait-Sahalia Y. and A.W. Lo. (2000). "Nonparametric Risk Management and Implied Risk Aversion." Journal of Econometrics 94, 9-51.

Bahra, B. (1997). "Implied Risk-Neutral Probability density Functions from Options Prices: Theory and Application." Working Paper, Bank of England.

Bakshi, G., N. Kapadia and D. Madan (2003). "Stock Return Characteristics, Skew Laws, and the Differential Pricing of Individual Equity Options." Review of Financial Studies 16, 101-143.

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

Bates, D. (1996). "Jumps and Stochastic Volatility: Exchange Rate Process Implicit in Deutsche Mark Options." Review of Financial Studies 9, 69-107.

Bliss, R. and N. Panigirtzoglou (2002). "Testing the Stability of Implied Probability Density Functions." Journal of Banking and Finance 26, 381-422.

Bliss, R. and N. Panigirtzoglou (2004). "Option Implied Risk Aversion Estimates." Journal of Finance 59, 407-446.

Bondarenko, O. (2003). "Estimation of Risk-Neutral Densities using Positive Convolution Approximation." Journal of Econometrics 116, 85-112.

Breeden, Douglas and Robert Litzenberger (1978). "Prices of State-Contingent Claims Implicit in Option Prices." Journal of Business 51, 621-652.

Bu, R. and K. Hadri (2007). "Estimating Option Implied risk-Neutral Densities using Spline and Hypergeometric Functions." Econometrics Journal 10, 216-244.

Campa, J.M., P.H.K. Chang and R.L. Reider (1998). "Implied Exchange Rate Distributions: Evidence from OTC Option Markets." Journal of International Money and Finance 17, 117-160.

Chicago Board Options Exchange (2003). VIX CBOE Volatility Index. http://www.cboe.com/micro/vix/vixwhite.pdf.

Danielsson, J. and Casper de Vries (2000). "Value-at-Risk and Extreme Returns." Working paper.

Derman, E. and I. Kani (1994). "Riding on a Smile." RISK 7 (Feb.), 32-39.

Derman, E. and I. Kani (1998). "Stochastic Implied Trees: Arbitrage Pricing with Stochastic Term and Strike Structure of Volatility." International Journal of Theoretical and Applied Finance 1, 61-110.

Dupire, B. (1994). "Pricing and Hedging with Smiles." RISK 7 (Jan), 18-20.

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

Embrechts, Paul, C. Klüppelberg, and T. Mikosch. 1997. Modelling Extremal Values for Insurance and Finance. Springer.

Gemmill, G. and A. Saflekos (2000). "How Useful are Implied Distributions? Evidence from Stock-Index Options." Journal of Derivatives 7, 83-98.

Jackwerth, J.C. (1997). "Generalized Binomial Trees." Journal of Derivatives 5 (Winter), 7-17.

Jackwerth, J.C. (1999). "Implied Binomial Trees: A Literature Review." Journal of Derivatives 7, 66-82.

Jackwerth, J.C. (2000). "Recovering Risk Aversion from Option Prices and Realized Returns." Review of Financial Studies 13, 433-451.

Jackwerth, J.C. (2004). Option-Implied Risk-Neutral Distributions and Risk Aversion. Charlotteville: Research Foundation of AIMR.

Jackwerth, J.C. and Mark Rubinstein (1996). "Recovering Probability Distributions from Option Prices." Journal of Finance 51, 1611-1631.

Jiang, G.J. and Y.S. Tian (2005). "The Model-Free Implied Volatility and Its Information Content." Review of Financial Studies 18, 1305-1342.

Malz, A.M. (1997). "Estimating the Probability Distribution of the Future Exchange Rate from Options Prices." Journal of Derivatives 5, 18-36.

McNeil, Alexander .J., Rudiger Frei, and Paul Embrechts. 2005. Quantitative Risk Management. Princeton University Press.

Melick W.R. and C.P. Thomas (1997). "Recovering an Asset's Implied PDF from Option Prices: An Application to Crude Oil during the Gulf Crisis." Journal of Financial and Quantitative Analysis 32, 91-115.

Optionmetrics (2003). "Ivy DB File and Data Reference Manual, Version 2.0." Downloadable .pdf file, available on WRDS.

Poon, S-H and C.W.J. Granger (2003). "Forecasting Volatility in Financial Markets: A Review." Journal of Economic Literature 41, 478-539.

Rosenberg, J. and R. Engle (2002). "Empirical Pricing Kernels." Journal of Financial Economics 64, 341-72.

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

Rubinstein, M. (1998). "Edgeworth Binomial Trees." Journal of Derivatives 5, 20-27.

Shimko, David. 1993. "The Bounds of Probability." RISK 6, 33-37.

Söderlind P. and L. Svensson (1997). "New Techniques to Extract Market Expectations from Financial Instruments." Journal of Monetary Economics 40, 383-429.

Table 1
S&P 500 Index Options Prices, Jan. 5, 2005

S&P 500 Index closing level, = 1183.74 **Interest rate = 2.69**
Option expiration: 3/18/2005 (72 days) **Dividend yield = 1.70**

	Calls				Puts			
Strike price	**Best bid**	**Best offer**	**Average price**	**Implied volatility**	**Best bid**	**Best offer**	**Average price**	**Implied volatility**
500	-	-	-	-	0.00	0.05	0.025	0.593
550	-	-	-	-	0.00	0.05	0.025	0.530
600	-	-	-	-	0.00	0.05	0.025	0.473
700	-	-	-	-	0.00	0.10	0.050	0.392
750	-	-	-	-	0.00	0.15	0.075	0.356
800	-	-	-	-	0.10	0.20	0.150	0.331
825	-	-	-	-	0.00	0.25	0.125	0.301
850	-	-	-	-	0.00	0.50	0.250	0.300
900	-	-	-	-	0.00	0.50	0.250	0.253
925	-	-	-	-	0.20	0.70	0.450	0.248
950	-	-	-	-	0.50	1.00	0.750	0.241
975	-	-	-	-	0.85	1.35	1.100	0.230
995	-	-	-	-	1.30	1.80	1.550	0.222
1005	-	-	-	-	1.50	2.00	1.750	0.217
1025	-	-	-	-	2.05	2.75	2.400	0.208
1050	134.50	136.50	135.500	0.118	3.00	3.50	3.250	0.193
1075	111.10	113.10	112.100	0.140	4.50	5.30	4.900	0.183
1100	88.60	90.60	89.600	0.143	6.80	7.80	7.300	0.172
1125	67.50	69.50	68.500	0.141	10.10	11.50	10.800	0.161
1150	48.20	50.20	49.200	0.135	15.60	17.20	16.400	0.152
1170	34.80	36.80	35.800	0.131	21.70	23.70	22.700	0.146
1175	31.50	33.50	32.500	0.129	23.50	25.50	24.500	0.144
1180	28.70	30.70	29.700	0.128	25.60	27.60	26.600	0.142
1190	23.30	25.30	24.300	0.126	30.30	32.30	31.300	0.141
1200	18.60	20.20	19.400	0.123	35.60	37.60	36.600	0.139
1205	16.60	18.20	17.400	0.123	38.40	40.40	39.400	0.139
1210	14.50	16.10	15.300	0.121	41.40	43.40	42.400	0.138
1215	12.90	14.50	13.700	0.122	44.60	46.60	45.600	0.138
1220	11.10	12.70	11.900	0.120	47.70	49.70	48.700	0.136
1225	9.90	10.90	10.400	0.119	51.40	53.40	52.400	0.137
1250	4.80	5.30	5.050	0.117	70.70	72.70	71.700	0.139
1275	1.80	2.30	2.050	0.114	92.80	94.80	93.800	0.147
1300	0.75	1.00	0.875	0.115	116.40	118.40	117.400	0.161
1325	0.10	0.60	0.350	0.116	140.80	142.80	141.800	0.179
1350	0.15	0.50	0.325	0.132	165.50	167.50	166.500	0.198
1400	0.00	0.50	0.250	0.157	-	-	-	-
1500	0.00	0.50	0.250	0.213	-	-	-	-

Source: Optionmetrics

Table 2: Summary Statistics on Fitted S&P 500 Risk Neutral Densities, Jan. 4, 1996 - February 20, 2008

	Average	Standard deviation	Minimum	Maximum
S&P Index	1140.60	234.75	598.48	1565.20
Days to expiration	54.2	23.6	14	94
Number of option prices				
# calls available	46.2	17.6	8	135
# calls used	37.6	15.4	7	107
IVs for calls used	0.262	0.180	0.061	3.101
# puts available	41.9	15.0	6	131
# puts used	32.9	12.4	6	114
IVs for puts used	0.238	0.100	0.062	1.339
Left tail				
α_{0L} connection point	0.8672	0.0546	0.6429	0.9678
ξ	0.0471	0.1864	-0.8941	0.9620
μ	1.0611	0.0969	0.9504	2.9588
σ	0.0735	0.0920	0.0020	2.2430
Right tail				
α_{0R} connection point	1.0900	0.0370	1.0211	1.2330
ξ	-0.1800	0.0707	-0.7248	0.0656
μ	1.0089	0.0085	0.8835	1.0596
σ	0.0416	0.0175	0.0114	0.2128

Note: Tail parameters refer to the risk neutral density expressed in terms of gross returns, S_T/S_0. "# calls (puts) available" is the number for which it was possible to compute implied volatilities. "# calls (puts) used" is the subset of those available that had bid prices of $0.50 and above.

Table 3: Summary Statistics on the Risk Neutral Density for Returns on the S&P 500, 1/4/1996 - 2/20/2008

	Mean	Std Dev	Quantile				
			.10	.25	.50	.75	.90
Expected return to expiration	0.61%	0.41%	0.13%	0.25%	0.52%	0.92%	1.22%
Expected return annualized	4.05%	1.89%	1.08%	2.03%	4.88%	5.46%	5.93%
Excess return relative to the riskless rate, annualized	-0.21%	0.43%	-0.57%	-0.30%	-0.16%	-0.04%	0.10%
Standard deviation	7.55%	2.86%	4.13%	5.49%	7.22%	9.34%	11.40%
Standard deviation annualized	20.10%	5.82%	12.80%	15.56%	19.67%	23.79%	27.57%
Skewness	-1.388	0.630	-2.165	-1.651	-1.291	-0.955	-0.730
Excess kurtosis	6.000	6.830	1.131	2.082	3.806	7.221	13.449
Skewness of RND on log returns	-2.353	1.289	-3.940	-2.834	-2.020	-1.508	-1.202
Excess kurtosis of RND on log returns	20.516	28.677	2.929	4.861	10.515	23.872	49.300

Note: The table summarizes properties of the Risk Neutral Densities fitted to market S&P 500 Index option prices, with GEV tails appended, as described in the text. The period covers 2761 days from 49 quarterly options expirations, with between 14 and 94 days to expiration. The RNDs are fitted in terms of gross return, S_T/S_0. Risk premium is the mean return, including dividends, under the RND minus LIBOR interpolated to match the time to expiration. Excess kurtosis is the kurtosis of the distribution minus 3.0. Skewness and Excess Kurtosis of RND on log returns are those moments from the fitted RNDs transformed to log returns, defined as $\log(S_T/S_0)$.

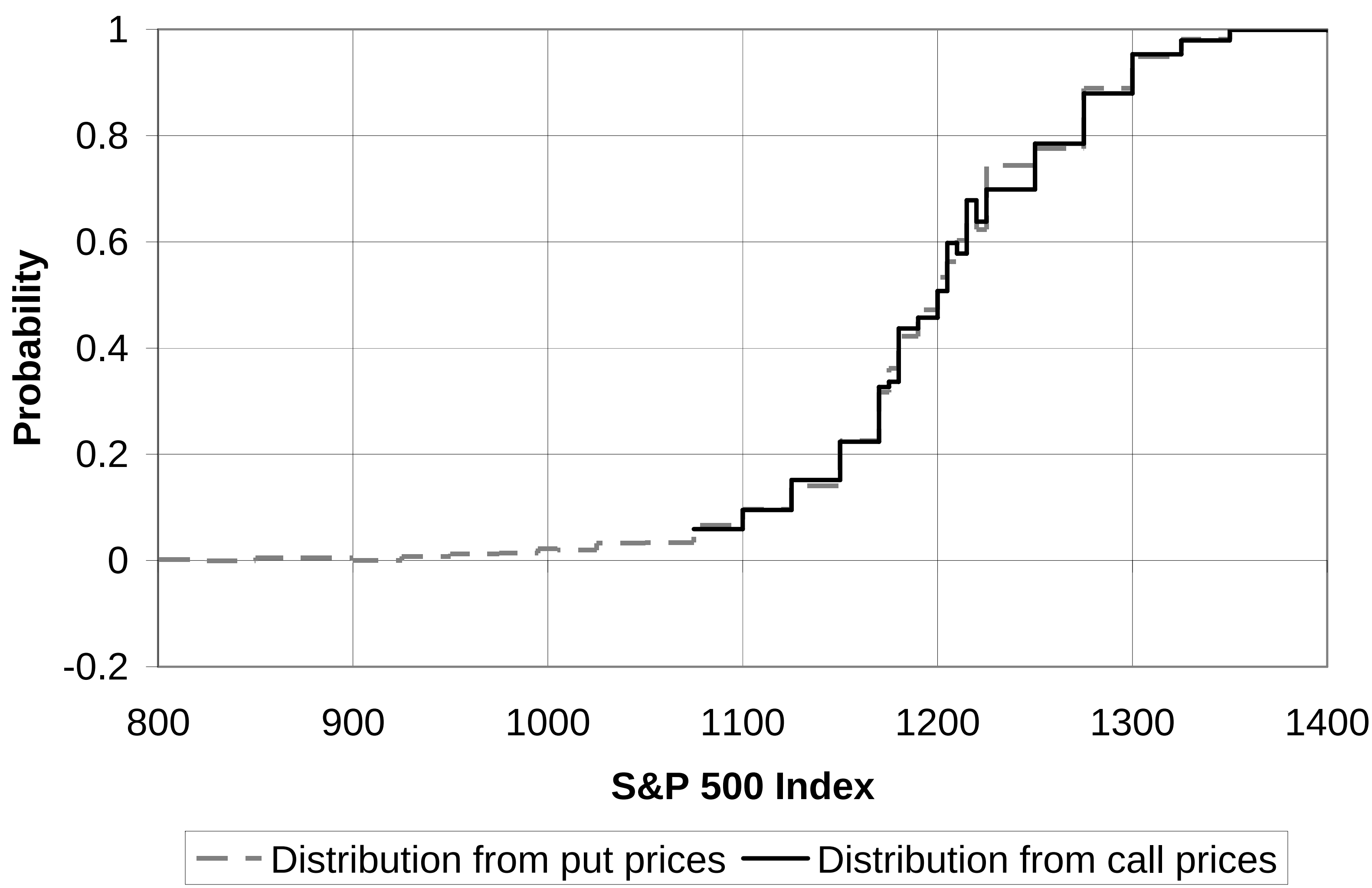
Table 4: Regression of Change in Quantile on Change in the Forward S&P Index Level

Regression equation: $\Delta RNDQ(t) = a + b\,\Delta F(t)$

The table shows the estimate b coefficient. t-statistics in parentheses

	Quantile										
	0.01	**0.02**	**0.05**	**0.10**	**0.25**	**0.50**	**0.75**	**0.90**	**0.95**	**0.98**	**0.99**
All observations	1.365	1.412	1.385	1.297	1.127	0.974	0.857	0.773	0.730	0.685	0.659
Nobs = 2712	(58.65)	(72.43)	(98.62)	(180.26)	(269.44)	(269.88)	(272.05)	(131.16)	(88.75)	(60.08)	(46.60)
Negative return	1.449	1.467	1.404	1.291	1.119	0.975	0.867	0.780	0.727	0.661	0.613
Nobs = 1298	(30.07)	(35.53)	(46.96)	(88.75)	(128.75)	(130.77)	(134.93)	(64.80)	(43.13)	(28.37)	(21.45)
Positive return	1.256	1.308	1.340	1.306	1.148	0.978	0.845	0.756	0.720	0.696	0.691
Nobs = 1414	(26.88)	(34.32)	(48.97)	(88.04)	(137.21)	(134.38)	(131.44)	(62.87)	(43.00)	(29.88)	(23.75)
Return < -1.0%	1.352	1.390	1.368	1.282	1.140	1.001	0.879	0.756	0.670	0.559	0.478
Nobs = 390	(12.64)	(14.26)	(19.62)	(39.18)	(54.80)	(59.48)	(61.05)	(27.71)	(17.94)	(11.28)	(8.12)
Return > 1.0%	1.106	1.194	1.292	1.310	1.173	0.988	0.843	0.756	0.726	0.710	0.712
Nobs = 395	(11.36)	(14.05)	(20.28)	(35.88)	(61.28)	(57.29)	(55.29)	(26.27)	(18.38)	(13.11)	(10.53)

Figure 1: Risk Neutral Distribution from Raw Options Prices



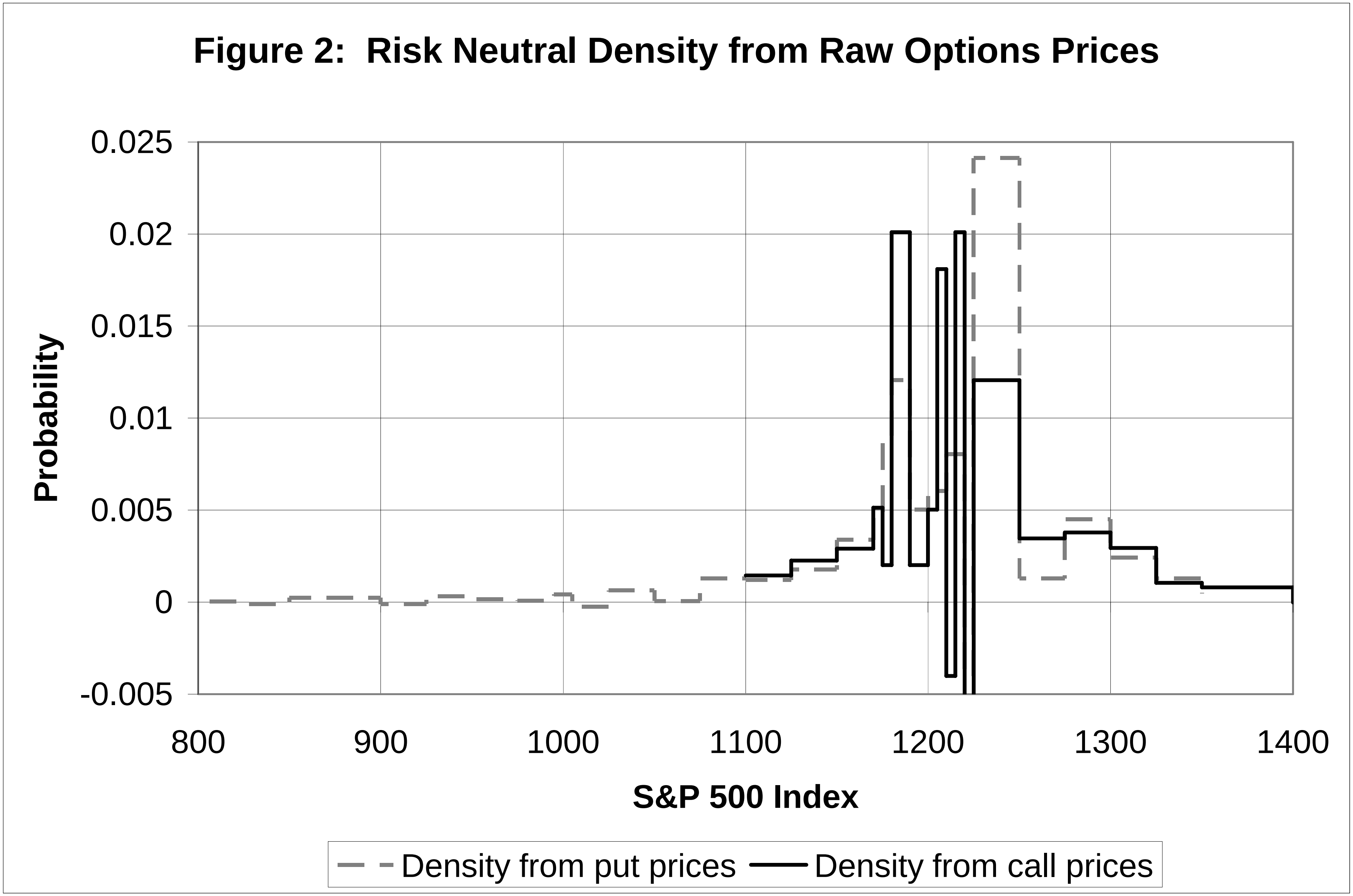


Figure 2: Risk Neutral Density from Raw Options Prices
Probability
0.025
0.02
0.015
0.01
0.005
0
-0.005
800
900
1000
1100
1200
1300
1400
S&P 500 Index
Density from put prices
Density from call prices

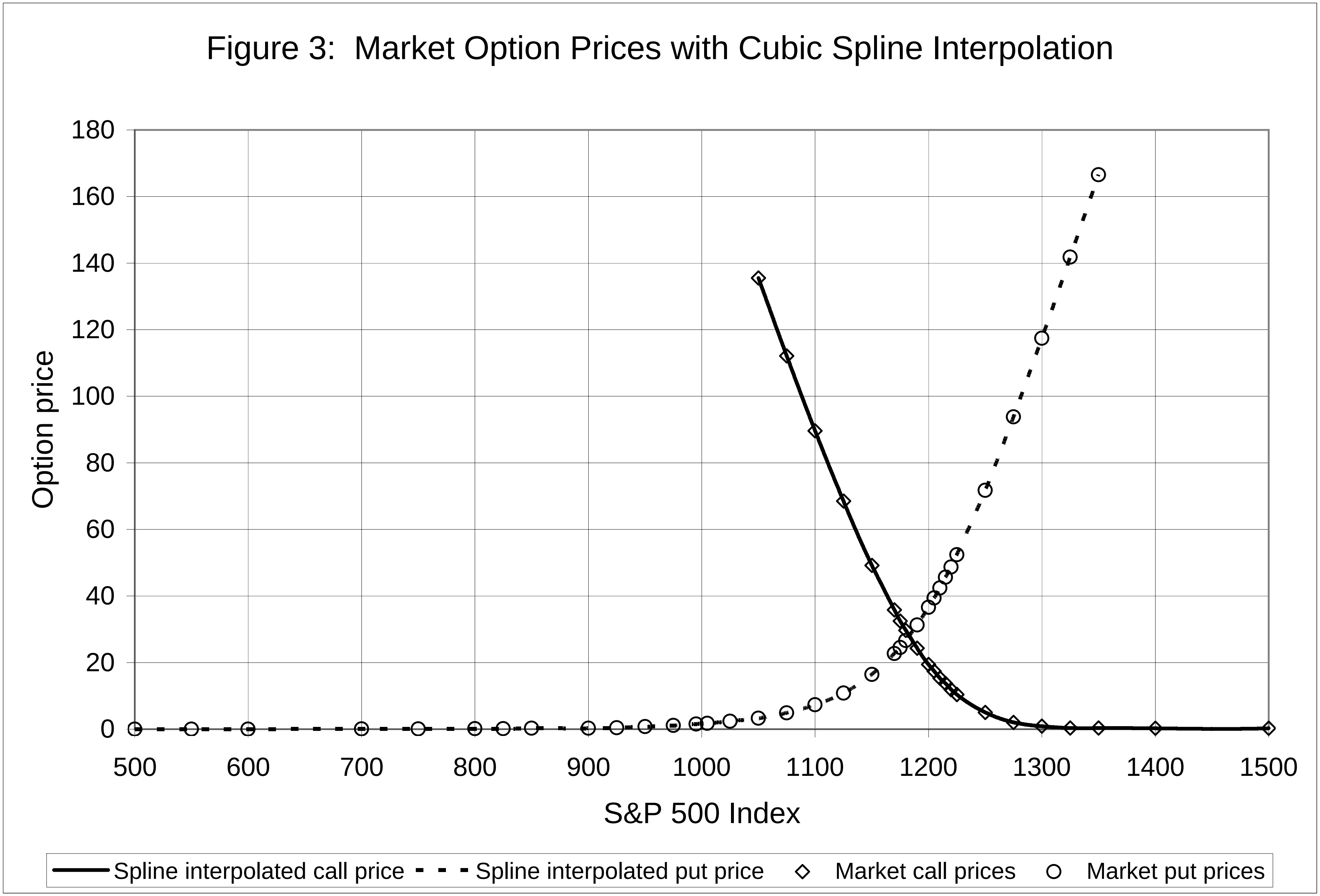


Figure 3: Market Option Prices with Cubic Spline Interpolation
180
160
140
120
100
80
60
40
20
0
500
600
700
800
900
1000
1100
1200
1300
1400
1500
Option price
S&P 500 Index
Spline interpolated call price
Spline interpolated put price
Market call prices
Market put prices

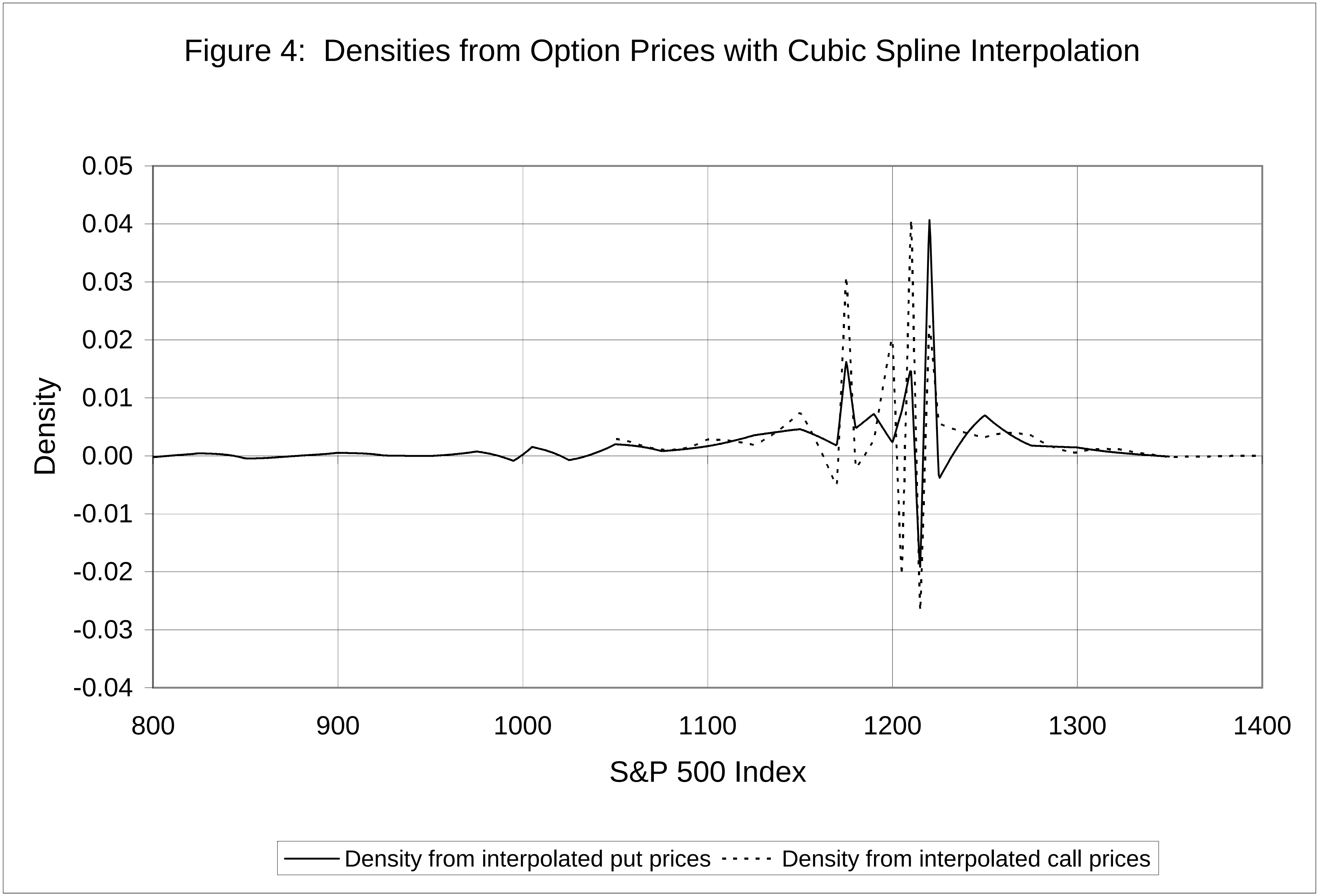


Figure 4: Densities from Option Prices with Cubic Spline Interpolation
Density
0.05
0.04
0.03
0.02
0.01
0.00
-0.01
-0.02
-0.03
-0.04
800
900
1000
1100
1200
1300
1400
S&P 500 Index
Density from interpolated put prices
Density from interpolated call prices

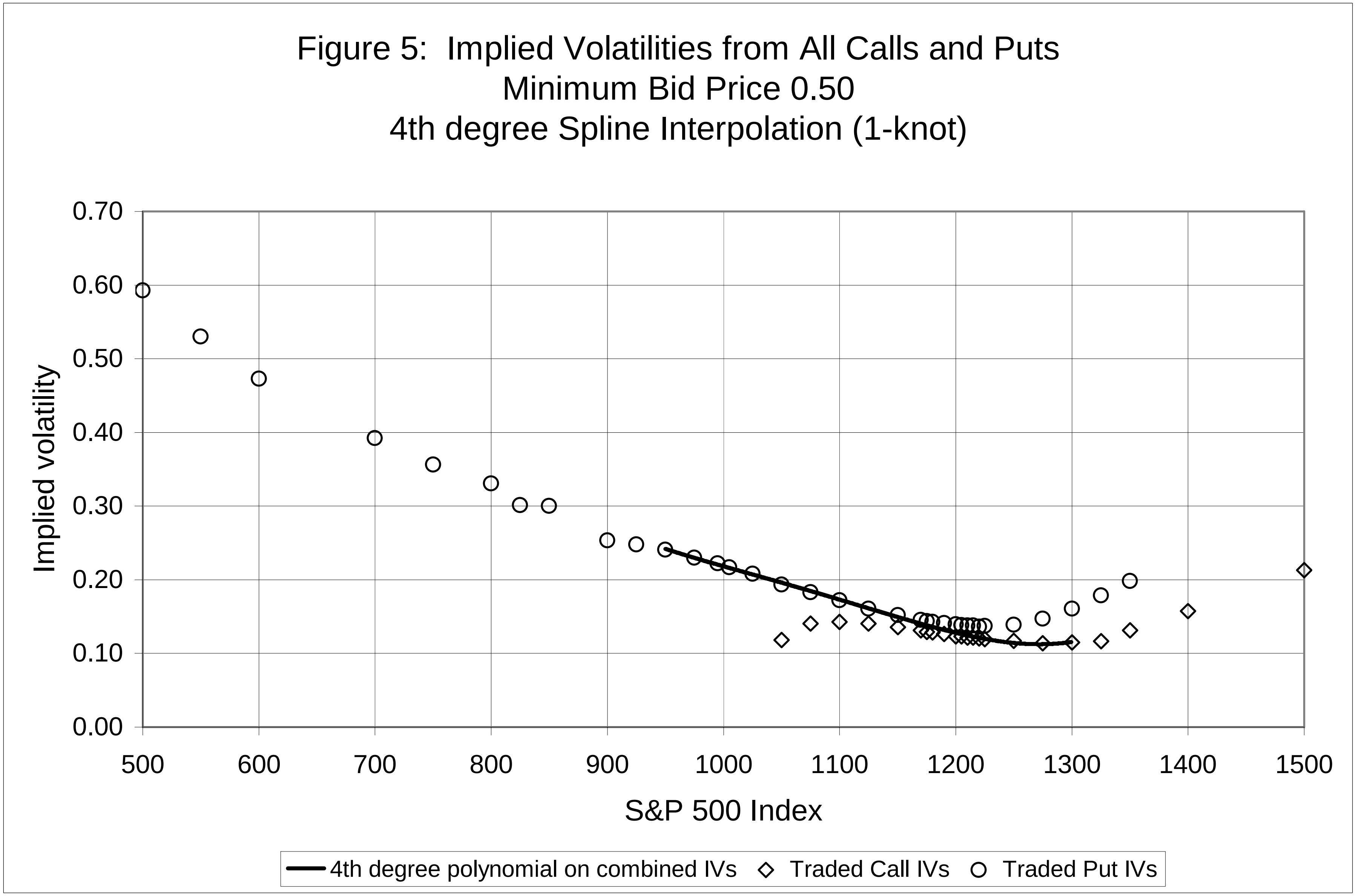


Figure 5: Implied Volatilities from All Calls and Puts
Minimum Bid Price 0.50
4th degree Spline Interpolation (1-knot)
Implied volatility
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
500
600
700
800
900
1000
1100
1200
1300
1400
1500
S&P 500 Index
4th degree polynomial on combined IVs
Traded Call IVs
Traded Put IVs

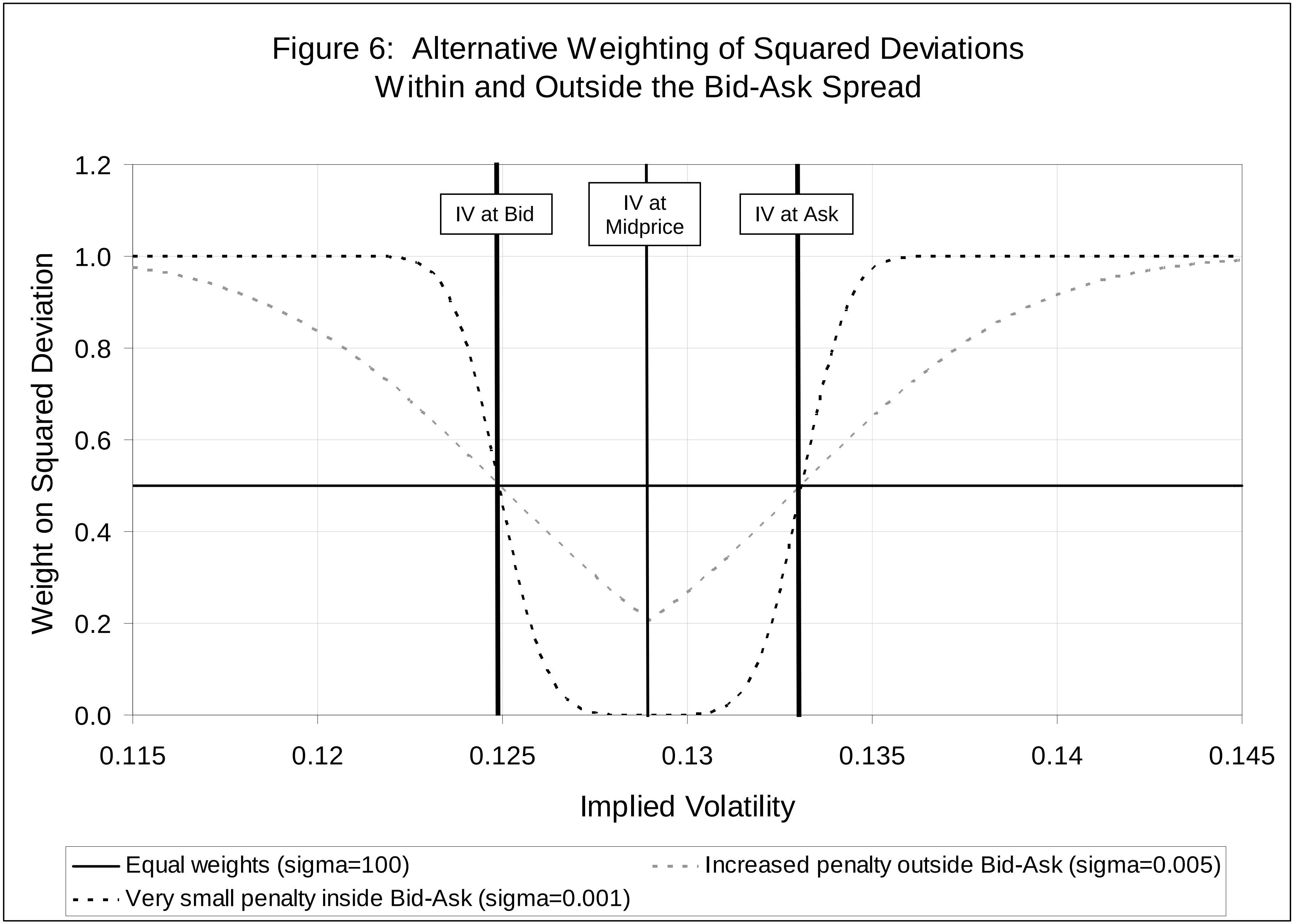


Figure 6: Alternative Weighting of Squared Deviations
Within and Outside the Bid-Ask Spread
IV at Bid
IV at Midprice
IV at Ask
1.2
1.0
0.8
0.6
0.4
0.2
0.0
Weight on Squared Deviation
0.115
0.12
0.125
0.13
0.135
0.14
0.145
Implied Volatility
Equal weights (sigma=100)
Increased penalty outside Bid-Ask (sigma=0.005)
Very small penalty inside Bid-Ask (sigma=0.001)

Figure 7: Densities Constructed Using Alternative Interpolation Methods

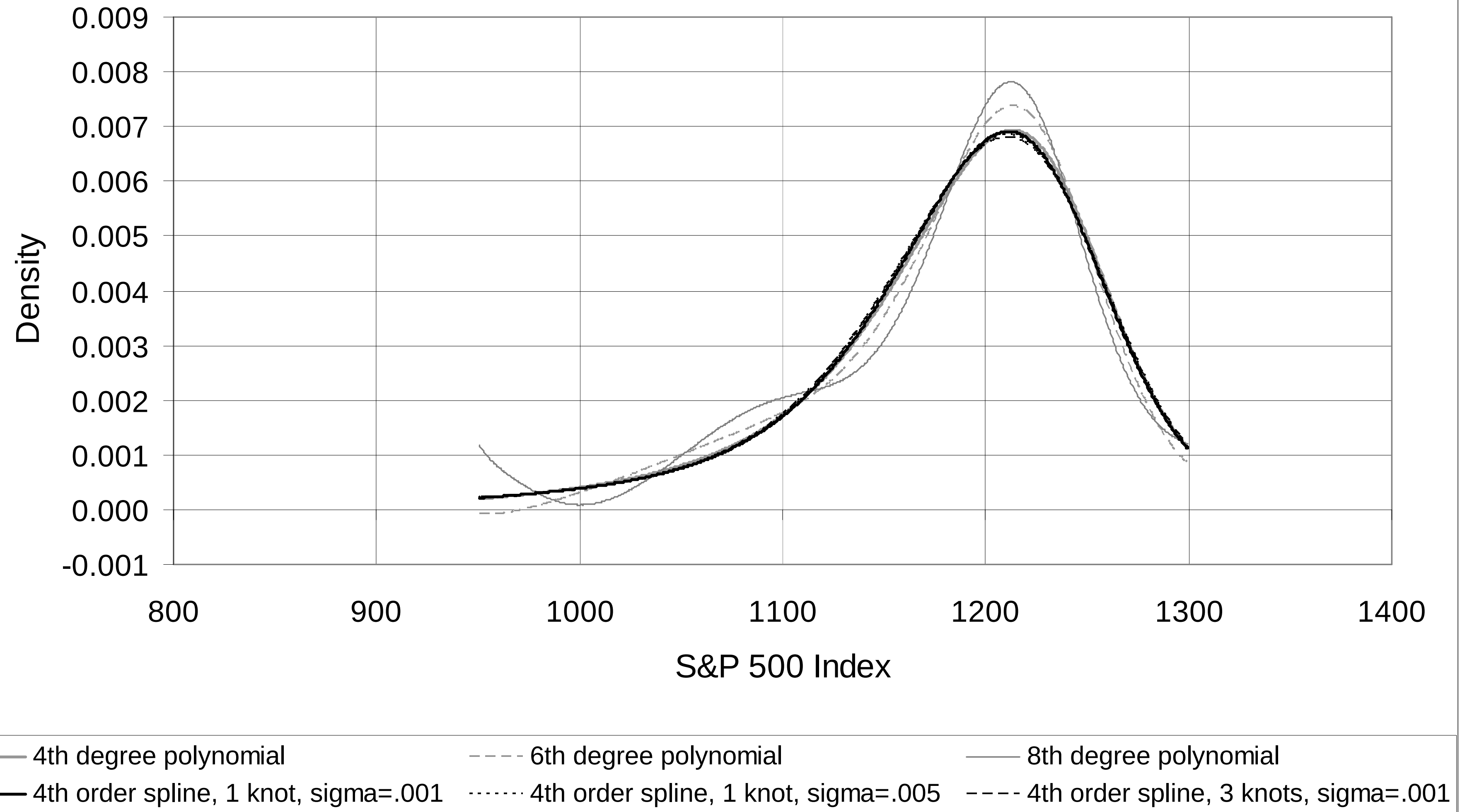


Figure 8: Risk Neutral Density and Fitted GEV Tail Functions

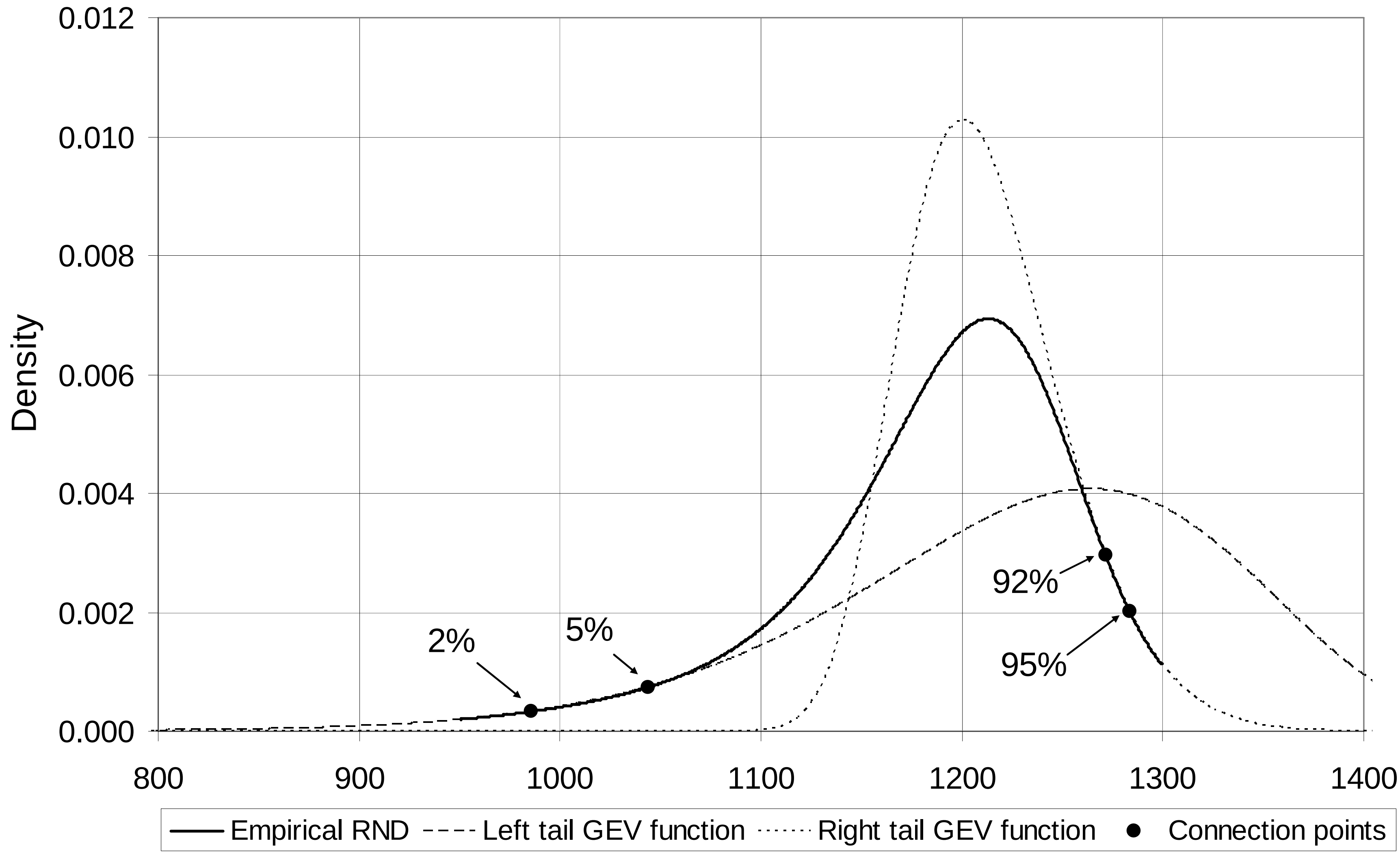


Figure 9: Full Estimated Risk Neutral Density Function for Jan. 5, 2005