Current draft: October 30, 2009
First draft: June 23, 2004

Information in Option Prices and the Underlying Asset Dynamics

Christopher G. Lamoureux
and
Alex Paseka[1]

We derive the exact joint transition density of returns and volatilities under Heston's (1993) stochastic volatility model. We reduce the computation of this density to the evaluation of a one-dimensional inverse Fourier transform, which allows us to use discretely-sampled data of any frequency and still avoid discretization bias in model estimation. We use weekly data on a broad panel of 15 options on the S&P 500 index and Markov Chain Monte Carlo methods. We show that parameter estimates are sensitive to the set of options data used to estimate the model. This data sensitivity reflects different skew and kurtosis patterns across moneyness and maturity categories. The volatility risk premium is instrumental in fitting the vertical position of a volatility smile without affecting its depth. We also confirm that tensions between the model and the data transcend discretization biases as well as parameter and state variable uncertainty.

JEL Classification: C90; D40; G10

Key Words: Joint Dynamics; Empirical Option Pricing; Bessel Bridge; Path Integral

[1]Department of Finance, The University of Arizona, Eller College of Business Administration, Tucson, 85721, 520–621–7488, lamoureu@lamfin.eller.arizona.edu; and paseka@cc.umanitoba.ca. We would like to thank Mike Chernov, Rob Indik, Yuriy Pinelis, Cameron Rookley, and Doug Witte.

We derive the (exact, discrete-time) joint transition density of returns and volatilities under Heston's (1993) stochastic volatility model to identify all model parameters and latent state variables. We use this derivation to add to the literature in four ways. First, we examine the effects of discretizing the continuous-time model on inference. Second, we construct exact predictive densities of the (Black and Scholes) implied volatility smile which we use for formal model inference. This produces our third contribution– a formal analysis of whether the documented difficulties of the model fitting the data remain after correcting for discretization, and the presence of parameter and state variable uncertainty. In particular, we examine whether the rich and varied patterns shown by implied volatilities in the data fall within model bounds in the presence of parameter and state variable uncertainty. Our fourth contribution is an analysis of the effect that different sets of options data have on estimation and inference. This sensitivity is reflected in practitioners' practice of estimating an option pricing model on different subsets of data, in an attempt to use an imperfect model to capture important features of each subset. While empirical studies have shown that the Heston model is not flexible enough to capture the patterns of smiles and smirks seen in the data, and researchers are working with a more general class of models that allows for jumps in both prices and volatilities, these extensions have not resolved all of the empirical issues with the simpler model.

Under Heston's model the evolution of stock returns under the actual probability measure P is:

$$dS_t = \mu S_t dt + \sqrt{v} S_t dz_t^P \tag{1}$$

$$dv_t = \kappa(\theta - v_t)dt + \sigma\sqrt{v_t} d\omega_t^P$$

The instantaneous correlation between the two Brownian motions (dz_t^P and $d\omega_t^P$) is ρdt. Appendix A reproduces Heston's solution to option prices $C(S, v, t)$ in this setting, and where the volatility risk premium is assumed proportional to the level of volatility (i.e., $\lambda \cdot v_t$).

Normally, since the exact joint transition density is recalcitrant, empirical research resorts to using Gaussian approximations - time-discretized versions of the stochastic differential

equations - of the joint transition densities to build likelihood functions.[1] Although sampling from time-discretized densities enjoys certain convergence properties under regularity conditions, it does introduce discretization bias into the parameter estimates. We start by demonstrating the different behavior of the Euler discretization from the actual continuous-time transition density. This difference is quite stark, especially in the tails, where the actual density looks more like the exponential density than its Gaussian (discretized) counterpart.

We use the information in returns and options simultaneously in estimation. It is well known that the stochastic volatility model used on returns data alone, has improper posteriors unless a proper prior is specified (see for example, Jacquier, Polson, and Rossi (2004), pp. 187-8). Our panel data has a relatively large number of options (15) and a relatively short time series (five years of weekly data). In this context, with only returns, the behavior of our likelihood is so poor that the posterior remains elusive unless one is willing to specify a very informative prior. Instead of doing this, we use options data to sharpen the likelihood. We use different subsets of options data and explore different weights on these data relative to the returns data in constructing posterior densities. This allows us to assess how the properties of the options data affect parameter estimates. Markov Chain Monte Carlo methods allow us to construct predictive densities of implied volatilities for our panels of options. Using these predictive densities, we examine the properties of the implied volatility smile under the model. These densities marginalize the parameter and state variable uncertainty that remains after using the data to estimate model parameters and state variables using the continuous-time model.

Structural econometric analysis of option pricing models is made complex since the model's absence of arbitrage restriction imposes a stark probabilistic error structure–residuals are strictly zero. In the context of Heston's (1993) model, given the parameters and the underlying asset price, on any particular observation date one option is all that is necessary to span the state space. One way to proceed in light of this stochastic singularity is to use the number of options equal to the number of unobserved state variables.[2] Information extracted

[1] See, e.g., Eraker, Johannes, and Polson (2003), Jones (2003), and Elerian, Chib and Shephard (2001). Chib, Pitt, and Shephard (2004) provide an augmentation algorithm for sampling from diffusions.

[2] If the number of assets is equal to the number of the states, it is possible, at least formally, to solve for the states exactly. Chen and Scott (1993) and Pearson and Sun (1994) use this method to obtain the exact

from a subset of options is then compared to another subset of options or the dynamics of the underlying asset. There are several drawbacks to the approach including the fact that it neglects potentially useful information in the options not used in the estimation; empirical comparisons tend to be ad-hoc–not the result of formal statistical inference, and the loss function used in estimation is often not properly aligned with inference (Christoffersen and Jacobs 2004). We use an alternative approach, in which the model price is equal to the observed price plus a measurement error.[3]

To explore different information in different option data sets, we report results for S&P500 index returns combined with four sets of options data, ranging from 15 options–comprising three distinct maturities, to a single option. With this panel we find that the variance-covariance matrix of pricing errors is (virtually) singular. In this case, the model is essentially unidentified, since all information comes from the option prices–the contribution of the underlying asset dynamics to the likelihood or posterior vanishes. We deal with this problem by setting all off-diagonal elements of the matrix to zero. This amounts to imposing a diagonal prior on this variance-covariance matrix.

We find that the average error in the model Black-Scholes implied standard deviation in the 15-option case is 1.21%. The average width of the 90%ile predictive posterior bands is .58%, and the width of the predictive posterior bands on the implied volatilities does not reflect the model fit. For example, the difference between the model and the data, in the metric of the implied standard deviation, is most severe in the shorter-term, deepest in-the-money call options. But the width of the predictive posterior band is average in this case. The widest predictive posterior band occurs for the deepest in-the-money, short-term option. These results are consistent with the observation that the risk-neutral density required to fit observed option prices exhibits more kurtosis than the Heston model can generate.

The significance of the volatility risk premium and its role in fitting volatility smiles are not clear in the existing literature. For example, Eraker (2004) finds that the premium is

likelihood, Aït-Sahalia and Kimmel (2007) to get an expansion of the likelihood, and Pan (2003) to obtain implied moments. For a more detailed list of applications see Garcia, Ghysels, and Renault (2003).

[3]Jacquier and Jarrow (2000) examine various ways to introduce errors into Black and Scholes option pricing model. For detailed descriptions see Johannes and Polson (2003) and Garcia, Ghysels, and Renault (2003).

statistically significant in a stochastic volatility model with correlated jumps in both returns and volatility yet insignificant in both a pure stochastic volatility model and a stochastic volatility model with jumps in returns only. Singleton (2006) points out that this difficulty in estimation may be the result of a limited history of options data and sometimes limited use of data on far out-of-the-money options. Our use of alternative option information sets allows us to better understand the role of this risk premium in option pricing.

We find that even though the volatility risk premium is significant across all information sets, its magnitude is highly sensitive to the information used to estimate the model. Its estimate becomes more precise as we move from fitting a single option to a smile or several smiles of different maturities. For example, when the information set consists of an intermediate term at-the-money option the posterior standard deviation of the volatility risk premium is 0.65, with a 5% posterior probability of this parameter being larger than -0.48. However, adding in-the-money and out-of-the-money intermediate term options to the at-the-money option lowers the standard deviation to 0.48 and the 5% posterior lower bound to -2.43. Our results suggest that parameters ρ and σ, affecting the skew and the depth of a smile, alone are insufficient to fit the model to the data, either under- or overpricing all options at the same time. The volatility risk premium provides an extra important degree of freedom to control the vertical location of the smile without changing its depth and skew.

We also find that fitting a smile at one maturity does not fit the smiles at other maturities. In particular, adding options to the information set that contain information about the intermediate maturity smile improves the model fit of short maturity in-the-money and long maturity out-of-the-money options even though these were not part of the estimation information set. However, in the same case the effect is opposite for short maturity out-of-the-money and long maturity in-the-money options. These properties result from two effects. With more options added to the information set posterior band narrows, and more data points fall outside the band. At the same time, higher kurtosis and more negative skew (high σ and low ρ)are more faithfully reproduced by the model.

These results suggest that combining options of different maturities may lead to relatively poor pricing of all options involved in the estimation. For example, using the prior to increase the pricing error variance on the single option used to estimate the model, only redistributes

pricing errors across different maturities with no clear overall improvement. In this case, the improved fit of implied volatilities on short-term and long-term options comes at the expense of those on intermediate term options.

The remainder of this paper is organized as follows. Section 1 contains a review of the empirical work done on option pricing models, with emphasis on the recent literature that estimates jointly the physical and risk-neutral processes from return dynamics and option prices. Section 2 provides a comparison between the actual transition density of volatility and its discretized counterpart. Section 3 presents the estimation approach used in this paper. Section 4 describes the data, which are unique to this study. Section 5 contains the results, and Section 6 concludes the paper. The relevant technical material is placed in appendices. Appendix A contains Heston's (1993) option pricing formula. Appendix B contains the derivation of the joint density of log prices and volatility, which is new to this paper. Appendix C shows the conditional densities used in our estimation.

1. Literature Review

The earliest tests of option pricing models are two-stage tests. As an example, Lamoureux and Lastrapes (1993) use a simple option pricing model (Hull and White 1987) to obtain stochastic implied volatilities from at-the-money options with a single expiration date. These are then compared to the realized variances using an encompassing regression. Even though the model used to price options is a no-arbitrage model, and therefore does not allow for a natural representation with pricing errors, the predicted variance from the model will have forecast errors. The model implies that the option-implied variances are the best, unbiased forecasts of future variance, so that any other information should be orthogonal to this forecast error. Lamoureux and Lastrapes document that while Black and Scholes implied volatilities of individual stock, intermediate term options have useful forecasting information, the forecast errors are not orthogonal to estimated historical variances. They infer from this that time-varying volatility is priced by the market.

Aït-Sahalia, Wang, and Yared (2001) use non-parametric methods to infer the state price density implied in a broad panel of three-month options using an empirical Girsanov change of measure. They also estimate the S&P 500 index diffusion process nonparametrically, and

then compare the densities. They conclude that the market does not price options efficiently under the null of a diffusion, and that adding the possibility of jumps (of magnitudes and numbers not experienced in the sample) is necessary to reconcile option prices to the underlying asset's dynamics.

Broadie, Chernov and Johannes (2005) examine options on S&P 500 futures over the period 1987 – 2003. They model jumps in the index and volatility, and also use a two-step approach. In fact, they take as given the parameters of the dynamic process estimated in Eraker, Johannes, and Polson (2003). These parameters are: $\theta = 0.023$, $\kappa = 5.8$, $\sigma = 0.35$, and $\rho = -0.40$. Broadie, Chernov and Johannes find strong evidence in support of jumps in prices with a jump risk premium of 2 – 5%.

More recently, empirical researchers have developed single-stage estimation and testing procedures which have advantages in efficiency over two-stage procedures. Most of these analyses use Heston's model as a base case. Identification requires information from both the return series of the underlying asset, which provides a glimpse of the physical measure, and option prices–determined in the equivalent martingale measure. In these procedures the option price itself is the focus of estimation, so an error is added to the pricing model. This error model is ad-hoc in the sense that the option pricing model is completely mute on its structure (Sims 2003).

Chernov and Ghysels (2000), Pan (2002), and Eraker (2004) all jointly estimate the dynamic process in the physical and risk-neutral measures but using different methods. Chernov and Ghysels use Heston's model to showcase the use of Semi-NonParametrics (SNP)/ Efficient Method of Moments (EMM) (as developed in Gallant and Tauchen (1996, 1998)). They use closing transaction prices on S&P 500 at-the-money, closest to maturity calls over the period November 1985 – October 1994. Their approach entails modeling the bivariate process: index returns and Black-Scholes implied volatilities. They use a SNP density - an approximation based on Hermite polynomial series expansion - to this end. They compute model parameters as follows: $\theta = .014$, $\kappa = 0.93$, $\lambda = -0.24$, $\rho = -0.018$, and $\sigma = 0.062$.[4]

Pan (2002) develops an implied-state GMM procedure. She obtains implied volatilities

[4]Here and in what follows all parameter estimates are reported according to parameterizations in equations (1) and (2), and are annualized. Note that the time scale of σ in (2) is $\frac{1}{\Delta t}$, not $\frac{1}{\sqrt{\Delta t}}$.

from near-the-money, short-term options conditional on the model parameters. She uses weekly data on S&P 500 options over the period January, 1989 through December, 1996. Her method allows her "to focus directly on the joint dynamics of the state variables rather than the market observables." (p. 11) Her parameter estimates are: $\theta = .0137$, $\kappa = 7.1$, $\lambda = -7.6$, $\rho = -0.53$, and $\sigma = 0.32$. She notes that the volatility risk premium "implies an explosive process under the 'risk-neutral' measure leading to severely overpriced long-dated options." (p. 3) To deal with this problem, Pan adds jumps to the dynamic process (and uses Bates' (1996) model of option pricing in the presence of jumps).

Eraker (2004) uses a random set of S&P 500 options from the period January 1987 through December, 1990. He uses a likelihood-based approach similar to that in this paper, but instead of working with the actual transition densities as we do, he uses a Gaussian approximation (essentially a Euler discretization).[5] Eraker specifies that the residuals from the option pricing model follow independent Ornstein-Uhlenbeck processes. His parameter estimates are: $\theta = 0.049$, $\kappa = 4.79$, $\lambda = -2.52$, $\rho = -0.57$, and $\sigma = 0.55$. Eraker finds that adding jumps to the Heston pricing model generates virtually no improvement in pricing.[6]

These three sets of estimates obtained from similar data are disparate. For example, the half-life of volatility shocks varies from 1.2 months to 1.7 months to 8.9 months. Similarly the volatility of market volatility varies from 6% to 55%. The steady state standard deviation of volatility ranges from 0.54% in Chernov and Ghysels to 1% per year in Pan. All of this suggests that a careful evaluation of the mapping between the model and the estimation methodology is in order.

2. Continuous vs. Discretized Transition Densities

We avoid discretization by using bivariate inverse Fourier transformation to obtain the exact solution for the joint transition density of the underlying index and its volatility.

[5]Broadie and Kaya (2004) show that discretization can induce a very large bias in option pricing when $\nu = \frac{2\kappa\theta}{\sigma^2} - 1$ is close to zero, and σ is large relative to θ. In this paper we are more interested in the impact of discretizing transition density on the parameter estimates.

[6]Jones (2003) also uses discretization and Gibbs sampling to estimate a stochastic volatility model. He finds that extensions to Heston's model cannot accommodate the empirically observed implied volatility smiles. Jones' emphasis is on using option prices to develop a more general model of stochastic volatility, as distinct from *testing* an option pricing model.

Normally, it would be computationally infeasible to use a bivariate inverse Fourier transform in estimation problems. However, we show that due to the affine structure of Heston's model it is possible to reduce the bivariate Fourier inversion to univariate inversion using an integral representation of Bessel functions. This reduction makes numerical computation of the joint density as tractable as that of Heston's option value. Thus, our approach is a deterministic quadrature method that exhibits exponential convergence (Judd (1998)) to arbitrary precision–a result which is independent of the available data frequency. By contrast, the Euler approximation based on Gaussian densities strongly depends on the time step (inverse of the data frequency). Broadie and Kaya (2004) show the deterioration in this method's reliability for option prices. In *estimation*, these poor results are aggravated by the lack of smoothness in the coefficients of the square root volatility process. By avoiding discretization bias, our approach makes it feasible to use lower frequency data than is required when the process is discretized. One advantage of this is that the use of lower frequency data, such as weekly, can mitigate the effects of trading and mechanism related noise that may perturb market data at higher frequencies.

Our estimation approach involves learning about v_t from all of the information in the sample, including r_t, r_{t-1}, and r_{t+1} along with v_{t-1} and v_{t+1}, and the entire vector of option prices at time t. This requires the *joint* transition density for returns and variance, which is derived in Appendix B.[7] Singleton (2001) shows the general approach of using Fourier representation or characteristic functions to derive these densities (see too, Revuz and Yor (1999), Geman and Yor (1993), and Feller (1951)). We show how to reduce the double integral (in equation 28) to a single integral (equation 39).

Table 1 shows the behavior of the transition density $f(\tilde{v}_{t+1}|\cdot)$ in the Euler discretized case (referred to as Method ED in the table) and in the continuous-time specification (referred to as Method GF in the table). This example uses the following parameters: $\mu = 0.08$, $\theta = 0.029$, $\kappa = 2.9$, $\sigma_v = 0.4$, $\rho = -.5$, $\tau = \frac{7}{365}$, $x_0 = \ln(500)$. The table reports properties of this conditional density including the value of the maximal density function and the value of

[7]Lamoureux and Witte (2002) also use exact (discrete-time) transition densities from the Feller process used by Cox, Ingersoll, and Ross (1985) to estimate a term structure model with orthogonal factors. The estimation approach is the same, but here we work with a bivariate process for returns and volatility.

$\tilde{v}_{t+1}$ at which this function is maximized (scaled to 1000 at their respective global maxima). Before turning to the comparison of the two methods, consider some of the properties of this conditional density. Note that the effect of ρ is important. In the case of a large negative return, (when P_{t+1} is 450), when v_t equals its unconditional mean, the median of the density of $\tilde{v}_{t+1}$ is .056. By contrast, for a positive return of similar magnitude (when P_{t+1} is 550), the analogous median of $\tilde{v}_{t+1}$ is .0213. Similarly, consider the case of 4.9% absolute weekly return (where P_{t+1} is either 475 or 525). If v_t is relatively high, at .048, then the median of the density of $\tilde{v}_{t+1}$ is .0576 for the negative return, and .038 for the positive return. Also note that in these cases, the conditional density of $\tilde{v}_{t+1}$ is much flatter than in the quiescent cases. When v_t is at its unconditional mean, in the presence of no change in the index value, the 98%ile range of $\tilde{v}_{t+1}$ is: [.013 .049]. In the case of the large positive return, with $v_t = .048$, this range is [.016 .059]. These densities often exhibit marked asymmetry. In the case where P_{t+1} equals P_t and v_t equals its unconditional mean, the maximum likelihood value of $\tilde{v}_{t+1}$ is .027, while the median of this density is .028. The 1%ile value is .015 less than the median, whereas the 99%ile value is .021 greater than the median.

This asymmetry contrasts with the discretized density. In the case where the return is 0, and v_t equals its unconditional mean, the discretized density is maximized at this mean value (.029), and the distribution is symmetric. Nevertheless, in quiescent periods the discretized density approximates the actual density reasonably. However, this approximation deteriorates in other settings. When the lagged volatility is small and the return is nonnegative, the discrete density includes negative values in its support. The actual density does not, but we see that it is also bounded above zero. Consider the case of the positive 4.9% return, when v_t is .025 (slightly below its unconditional mean). The interquartile range of $\tilde{v}_{t+1}$ is [.009 .031] while under the discrete approximation this is [.003 .028]. Here again the actual density exhibits marked skewness that is absent from the discrete approximation: the maximum likelihood value of $\tilde{v}_{t+1}$ is .017, while the median is .018. The Euler approximation has both at .015. Since stochastic volatility models are designed to deal with leptokurtic patterns in the data, this degenerate tail behavior of the discretized version may have significant effects on parameter and volatility estimates, with adverse consequences for estimating option prices.

There are two metrics to judge the approximation error of a discretization algorithm: 1) strong convergence which measures the mean square distance of the generated path from its exact continuous counterpart, and 2) weak convergence which measures the convergence of mean values (i.e., convergence in distribution). Kloeden and Platen (1999) show that the Euler scheme has strong order of convergence 1/2. The convergence property requires among other things that the drift and diffusion coefficients be Lipschitz-continuous. The square root volatility process in Heston's model does not satisfy this requirement. It becomes important to assess the impact of the breach of this requirement on parameter estimates.

Evidence presented by Broadie and Kaya (2004) concerns mostly the convergence in distribution properties of the Euler algorithm - the impact of the discretization bias on expectations like option prices. They find that although Euler method is faster, it becomes suboptimal when either the precision of the estimates is important or the parameter $\nu = \frac{2\kappa\theta}{\sigma^2} - 1$ is close to zero. In this paper we are concerned with strong convergence; i.e., the approximation errors in the transition density and how parameter estimates are affected by these errors.

3. Estimation

In addition to remaining faithful to the model cast in continuous time, we want to be sensitive to parameter and state variable uncertainty by integrating over the relevant spaces–rather than plugging in (consistent) estimators, when evaluating marginal posterior densities. Of the set of papers reviewed above that pursue joint estimation, we are most like Eraker (2004), in that we use Markov Chain Monte Carlo (MCMC) methods to integrate over all unknown spaces in estimation. However, we do not discretize. Since we work with the exact bivariate transition density for returns and variance, it is not possible to use a Gibbs sampler (as in Eraker). Instead, we use a Metropolis-Hastings algorithm described below.[8] Our approach is also like Aït-Sahalia, Wang, and Yared (2001), in that we want to use a broad cross-section of options–with different strikes and terms–to learn about model parameters.

Our MCMC approach draws each parameter and state variable in turn from its full

[8]Chib and Greenberg (1995) is a classic reference on the Metropolis algorithm. See also Robert and Casella (2004), pp. 267–320.

conditional density (taking all other parameters and state variables as given). This enables construction of marginal densities of the parameters and latent state variables, as well as functions of interest of these objects. These posteriors are integrated over the uncertainty in this space of parameters and state variables that remains after updating prior beliefs using all of the information contained in the data. The remainder of this section describes the MCMC method in more detail. For all analysis here, an improper diffuse prior is used, so that the posterior densities characterize the likelihood function. All full conditional densities for parameters and volatilities are described in Appendix C.

3a. Drawing v_t

The complexity in this transition density arises because the underlying model is cast in continuous time while the data are observed at discrete points. In the discrete-time model of Jacquier, Polson, and Rossi (1994), for example, the full conditional density for the volatility at time t is the product of three densities: the (univariate) transition densities of volatility from $t-1$ to t, and from t to $t+1$; along with the (univariate) conditional density of the return at time t on contemporaneous volatility. The difference between the continuous specification here and the discrete-time model comes not from the volatility per se–for which a discrete-time transition density exists–but from the density of the asset return conditional on the volatility process. The transition from x_t (the log of the price plus accumulated dividends) to x_{t+1} depends on the *path* of volatility over that interval.

The Metropolis algorithm requires a proposal density $q(\cdot)$, from which a candidate parameter or state variable is drawn. The current (x) and proposed (y) draws are compared: $\varrho = min\left\{\frac{f(y)}{f(x)}\frac{q(x|y)}{q(y|x)}, 1\right\}$. ϱ is compared to a random uniform (0,1) draw, u, and y replaces x if $\varrho > u$. Otherwise, the old draw, x, is kept. In this setting, there are no obvious proposal densities for v, so we use a Newton-Raphson procedure to maximize the full conditional density as a function of v_t. The resulting "maximum likelihood" estimate of v_t then serves as the mode, and the negative inverse Hessian, the variance of the normal kernel that serves as the proposal density at this step[9]. This Newton-Raphson procedure is implemented every K draws. So that intervening draws use the same proposal density. For all results reported

[9]The early application of the method is due to Chib and Greenberg (1994). Among more recent papers using the approach is Elerian et al. (2001).

here, we set $K = 30$.

Initial and terminal states are also drawn in each step of the MCMC: v_0 is drawn conditional on (the actual) x_0, x_1, and the on-hand draw from v_1; and v_{T+1} is drawn conditional on (the actual) x_{T+1}, x_T, and the on-hand draw from v_T, as described in Appendix C.

3b. Drawing Σ

Conditional on $\{v\}$, all other parameters and the data, the residual variance-covariance matrix has a standard inverse-Wishart form (that follows from the assumption that the pricing errors are multivariate normal). Thus: $\Sigma^{-1} \sim W(T, \sum_{t=1}^{T} \epsilon_t \epsilon_t')$. Conditional on model parameters and the entire vector of state variables ($\{v\}$), the ϵ_t (a 15-vector in our base case) are directly observed. In general, draws from this multivariate density are made by recognizing that the Choleski factorization may be obtained by taking univariate draws from a χ^2 density for the diagonal elements and a univariate normal density for the off-diagonal elements (see Muirhead (1982), p. 99). We draw Σ using a Gibbs sampling approach.

3c. Drawing μ

The mean parameter, μ is not identified in the equivalent risk-neutral measure, so we learn about μ (only) from the joint transition density (Appendix C). Again, given the complexity of this function, we design a Metropolis step to draw μ. For the proposal density, we choose a Student t distribution with 14 degrees of freedom (i.e., $\nu = 14$). The mean of the proposal distribution is the mean of the log-returns in the sample ($\overline{x}$). The variance is the sample variance of the realized log-returns (s). This leads to a log-proposal kernel:

$$q(\mu) = -\frac{1}{2}(\nu + 1) \cdot \ln\left(1 + \frac{(\mu - \overline{x})^2}{\nu \cdot s}\right) \tag{2}$$

This proposal kernel is constant over the entire chain.

3d. Drawing θ, ρ, κ, and σ

All of these parameters appear in both the option prices and the joint transition density. In light of the complexity of these functions we also use the Newton-Raphson algorithm, as with the v draws, to get the mode and variance of the proposal normal kernels. So each K draws we obtain the maximum likelihood estimate (i.e., the mode of the posterior) and

variance, which are used as the mean and variance of a normal proposal density for the next K draws. (In the analysis below, K is set to 30 for all parameters.)

Well-behaved transition densities require that the following condition always hold: $\frac{2\kappa\theta}{\sigma^2} > 1$. We impose this condition, which ensures that volatility is strictly positive, on every draw of each of these parameters, using an accept-reject method from the *proposal* density.[10]

3e. Drawing λ

The parameter λ only appears in the option prices. Even so, the complexity of this density leads us to use the two step Newton-Raphson algorithm to obtain the first two moments of the normal proposal kernel. As with the other parameters, the proposal density is updated every K draws, and in the sequel K is set to 30.

4. Data

We obtain weekly quotes on S&P 500 call options from the Berkeley Options Data Base. The period is January 3, 1989 through October 6, 1993 (249 weeks). The option price is defined as the mid-point of the bid-ask spread at 9:00 am CST on every Tuesday in 1989, Wednesday in 1990, 1992, and 1993, and Thursday in 1991. The sample contains seven options with a near expiration (of no less than seven days), five options with the second-nearest expiration, and three options from the third-nearest expiration. The contemporaneous spot price on the S&P 500 Index is taken from the Berkeley tape for each option.[11] On four dates the most out-of-the-money, closest-to-expiration calls were not updated, following an opening quote of 0 bid and either 6.25 or 12.5 cents asked at 8:30. In these cases we use the midpoint of the opening quote.[12] These 15 options that comprise our sample are only a small portion of the S&P 500 complex. For example, on January 17, 1989, the index was around 284. On this date, 71 different call options were quoted. The available expiration months were: January, February, March, June, September, and December. Strike prices ranged from 175 to 325. Our sample contains February options with strike prices ranging from 270 to 300

[10]The ability to impose this constraint universally is clearly an advantage of the Bayesian approach.

[11]This is because the data are irregularly updated. So while 9:00 am is our target time, on a given day we may have option quotes from 8:58 through 9:02.

[12]Interestingly, there is trade activity at the quoted ask price throughout the day in these options.

in $5 increments; March options with strike prices ranging from 275 to 295 in $5 increments; and June options with strike prices of 280, 285 and 290. SPX options are European-style. We adjust for the dividend by subtracting the present value of the actual dividends paid over the option's remaining life from the index level, for the purposes of option valuation.

Rates on T-Bills whose terms differ by 1 day from each option's are hand collected from the *Wall Street Journal*. The mid-point of the bid and ask discount quotes is used to compute the yield. Dividend data are obtained from the S&P 500 Daily Dividend Record.

Properties of the sample options are displayed in Table 2. The average term of the seven short-term options (numbered 1 through 7) is about 22 days, with a minimum of 9 days and a maximum of 39 days. The term of the intermediate options (options 8 through 12) averages 56 days, with minimum of 37 days, and maximum of 129 days. The longest term options (Options 13 through 15) average 112 days, with a minimum of 65 days, and maximum of 220 days. The deepest in-the-money option is option 1, where the index level is on average 5% higher than the strike price. It is clear that options 1 and 15 exhibit the widest variation in moneyness over the sample. This is because we seek actively traded options that conform to the monotonicity, and the deep in-the-money calls with short terms to maturity are illiquid. The same holds for the deep out-of-the-money calls with long terms to maturity. The sample exhibits the familiar smirk patterns of implied standard deviations from the Black and Scholes model (isds).[13] For all three terms, the average isd increases monotonically in moneyness. By the same token, the variation in isds over the sample also increases monotonically in moneyness. Comparing the at-the-money options across terms (i.e., options 4, 10, and 14), we see a monotonic increase in the implied volatilities in term from 15.1% to 16.3%. The dispersion in the implied volatilities over this period is also large. The shortest-term at-the-money option's implied volatility ranged from a low of 9.1% (on January 17, 1991–week 107 in the sample) to a high of 29.4% (on February 3, 1993–week 214 in the sample). The dividends paid over the option lifetimes in the sample range from

[13]Black and Scholes implied volatilities are widely used by traders as a summary measure of an option's value. For example, a deep out-of-the-money call with a price of $0.125 may be "expensive" with an implied volatility of 26%, while a deep in-the-money call with a price of $25 may be "cheap" with an implied volatility of 15%. The use of these does not mean that we take the Black and Scholes model to be true, or even that we are using it as a benchmark.

an average of \$0.68 for the shortest term, to \$1.84 for the intermediate term, and \$3.69 for the longest term options. Since dividend payments are lumpy, (and the option terms vary as noted above), there is quite a large variation in dividend payments over the sample.

5. Results

5a. Different sets of Options

As noted in the introduction, with our sample of 249 weekly observations, the likelihood in the case of returns-only, is very ill-behaved. Very tight priors are needed to keep the sampler out of regions of joint parameter/$\{v\}$ space that are degenerate, and absorbing. To deal with this we add option price data to the returns on the index. In this first section of our results, we estimate the model using index returns combined with four sets of options data: 1) All 15 options in the sample; 2) Options 8, 10, and 12 (the deepest-in-, at-, and deepest out-of –the money options with intermediate term; 3) Option 10, with a diffuse prior on Σ; and 4) Option 10, with a sharp prior that $\Sigma_{1,1} = 0.2$ (note that in these last two cases Σ is a scalar). This fourth case puts the least information on the option prices, as the imposed error variance of 0.2 is 7.4 times larger than the mean of the posterior on this parameter from the single option, diffuse prior case. Our goal is to characterize how the posterior changes with different sets of options data.

In the cases with more than one option, when we impose a diffuse prior on all parameters, estimation degenerates in that Σ is virtually singular. Since the prior specifies that Σ must be positive semi-definite, each Gibbs draw (described in Appendix C) is technically of full rank, but if we only evaluate the first twelve significant digits, the matrix is singular. Table 3 shows a late draw from Σ from the 15-option case that manifests this problem. The diagonal elements are (pricing) error standard deviations, expressed in \$US, and the off-diagonal elements are correlations. We see that many of the pairwise correlations are greater than 96%. In general, the highest correlations are between options that expire on the same date. Since we started with a diffuse prior, one interpretation of these high correlations is that both moneyness and term affect option prices in manners not captured by the model. This

is explored below, when we look at the implied volatility smiles in detail.[14]

A natural solution to this problem is to remove some of the options from the sample. However, the focus in this paper is extracting information from both options and returns data so we take a different tack. If the pricing model held, pricing errors at a point in time would be uncorrelated across the different options. Therefore, we assume that covariance matrix of 15 option pricing errors is diagonal, that is we specify a diagonal prior on Σ. This was also done by Schotman (1996) in a study of Dutch bond and bond option prices, using a (one-factor) Vasicek model. This is an important use of a prior, since a singular Σ marginalizes the information in the time series of the underlying index. In this case, the model is not identified. Thus it is clear that this specification of the prior places more weight on the information on the underlying asset's return time-series relative to the cross-section of option prices–facilitating identification.

In all cases, we use 200,000 draws as a burn-in, and then keep the next 100,000 draws. The Metropolis algorithm either keeps the old value or accepts the proposed value at each draw. We also monitor acceptance probabilities in an effort to make the algorithm traverse all regions of the parameter space. Excessively high acceptance rates would imply that the algorithm moves slowly in parameter space. Low acceptance rates would mean that too many draws are wasted inefficiently due to poor choice of the proposal. The acceptance rates for each of the parameters obtained using the Metropolis algorithm average as follows: θ 60%, κ 58%, σ 49%, ρ 49%, μ 59%, and λ 54%.

Table 4 contains aspects of the posterior density of the six model parameters in the physical measure, along with the order of the Bessel function plus one $\left(\frac{2\kappa\theta}{\sigma^2}\right)$, and the steady-state variance of the stock return variance, $\sigma^2_{ss} = \left(\frac{\sigma^2\theta}{2\kappa}\right)$, for all four cases. In addition, we report the equivalent martingale measure parameters $\kappa^Q = \kappa + \lambda$ and $\theta^Q = \frac{\kappa\theta}{\kappa+\lambda}$.

The parameter estimates are sensitive to the information set. The constraint that the order of the Bessel function, $(\nu = \frac{2\kappa\theta}{\sigma^2} - 1)$ be positive is binding in the 15 options case, and perhaps as well in the 3 options case. It is not binding in the two cases with the single

[14]These high error correlations could be used to infer that the model does not fit the data. But from a Bayesian perspective, we know the model is wrong. Our interests include the extent to which it provides a useful framework for organizing the information in the data.

option. While κ is centered around 3 in both the 15 Option and the one option–constrained Σ case, this has a posterior mean of 8.9 from the three option case and 4.0 from one option–unconstrained Σ case. The dispersion of the posterior on κ is smallest in the 15 option case.

Our ρ estimate in the 15 option case is smaller than previous results in the literature, but the value is closer to zero in the other three cases. The unusually low value of ρ in the 15 option case is perhaps the result of strong skew imposed by the short term volatility smile in the data. The largest value of ρ is seen in the case of one option–unconstrained Σ, where the mean of the posterior is -.63. Not surprisingly, we see that having more options in the sample tightens the posterior on λ. In the two single option cases, the 90%ile posterior ranges of λ are $[-2.7 \ -0.5]$ (when Σ is unconstrained), and $[-2.6 \ -0.4]$, (in the Σ constrained case). The two cases with only a single option contain no direct empirical information about the smile. We see the smallest values of σ_v in these two cases. Further, σ_v is smaller when we use the prior to specify a larger variance on the pricing error. This is consistent with the tension between the kurtosis in the risk-neutral density implied by the options prices and that observed in the data. This confirms the intuition that the smile implies leptokurtosis in the (risk-neutral) density of the underlying asset. Moving from 15 to 3 options, (allowing the pricing error variance to shrink), the parameters adapt to more faithfully price the options. In this case σ_v is therefore larger as there is no longer a constraint imposed by the relative flatness of the long term smile (options 13,14, and 15). To bring the skewness in line, ρ is also smaller in absolute value in this case since (following Das and Sundaram (1999)) ρ has a first order effect on skewness and only a second-order effect on kurtosis. A simple computation reveals that with λ unchanged at its 15-option case level, the unwanted side effect of changing σ and ρ, however, would be the downward shift of the entire smile, leading to underpricing of all intermediate term options. Therefore, λ adjusts downward (from -1.6 to -3.1) to raise the smile and reverse the underpricing.

As noted in the introduction, researchers and practitioners have focused on models' abilities to fit observed Black-Scholes implied volatility patterns. We are in a position to formally compare the observed isds with those from the estimated model. For each option in the sample, we construct a predictive posterior density of the isd, from each of the four cases. We do

this by defining the option price under the model and then obtaining the isd from that price, for each of the 100,000 Metropolis draws. These draws on the isds comprise the marginal predictive density of the isd that integrates over the unknown parameters and that day's volatility. Table 5 provides root-mean-square errors (RMSEs), predictive posterior band widths, and the relationships between the actual isds and these bands for each of the 15 options in the sample for each of the four information sets. So for example, for the at-the-money, intermediate term option (Option 10), the square-root of the average of the squared error between the realized isd on week t and the mean of the predictive density of that isd, (averaged over the 249 weeks in the sample) is 0.333%, when all fifteen options are used in estimation. In this case, the 90%ile bands have an average width of 0.561%. Of the 249 weeks, in 35 (14%) the actual isd lies above the 90%ile band and in 42 (16%) the actual isd lies below this band.

Adding options 8 and 12 to the information set (relative to only option 10), places more emphasis on kurtosis. This puts upward pressure on σ_v and transfers more weight to the information in the option data relative to that in returns. In this case we are fitting the smile represented by these three options. Kurtosis implied by this smile is much higher than that implied by returns, which leads to a diminished role for the returns data on the estimates. The standard deviation of price error for option 10 drops from 16.4% to 4% as we add options 8 and 12 to the information set. This drop in the pricing error translates into a decline in the mean RMSE for option 10's implied volatility from 0.315% to 0.008%.

By far the best overall fit in Table 5 is for option 10 (the intermediate-term at-the-money option) in the case where the information set comprises options 8, 10, and 12. In this case, the fit is almost perfect. The 90%ile posterior band on the implied volatility is 26 basis points (0.26%) wide, and all 249 volatilities lie within this band. This happens because of the effect that the data have on λ, seen also in κ^Q, and σ. When confronted with the information in the (intermediate-term) smile, σ increases dramatically to match the depth of the empirical smile. While in-the-money option prices remain essentially unchanged, both at- and out-of-the-money option prices decline, with the latter prices declining the most. With the negative ρ this also results in a stronger skew: the increase in kurtosis being largely allocated to the left tail, the reason for the large decline in out-of-the-money option prices. We see the effect

that the empirical smile has on λ in the equivalent martingale measure parameters, θ^Q and κ^Q. κ^Q increases by a factor of 2.5× to accommodate the smile. By contrast, when we add information from smiles at longer and shorter maturities, λ and κ^Q revert to their levels with a single option. So while this model can fit a smile, it cannot fit smiles at various maturities. An extension that this suggests might be a stochastic volatility risk premium that is correlated with the return process and is a function of option maturity.

As we look across moneyness categories in Table 5, going from the use of a single option to 3 options, the RMSEs for at-the-money options that were not used to estimate parameters and volatilities deteriorate. Both options 4 and 14 RMSEs are larger in the single option case. At the same time, at the short maturity the RMSEs of the in-the-money options shrink whereas those of the out-of-the-money options increase, when we add the two intermediate term options to the information set. For the long maturity, the effect is opposite. Adding options 8 and 12 increases RMSE of in-the-money option 13 and reduces RMSE of out-of-the-money option 15.

As can be seen in Table 5, the best fit between the model and the data when all fifteen options are used in estimation, is for option 11, the closest out-of-the-money, intermediate term option. The worst fit is the short term, deepest in-the-money call (option 1). The root-mean-square error over the 249 sample weeks is 3.6%, relative to a 90%ile band width of 0.57%. Since 193 of the 249 isds lie above the band, there is a clear bias–the model underprices these options (which are equivalent through put-call-parity to short-term deep out-of-the-money puts).

5b. The Smile and all 15 options

In this next set of results, we examine the actual smile and the posterior from all of the information in more detail. Figures 1 through 3 provide a graphical depiction of the relationship between the actual and model isds for the 15 option case. All of the figures show the predictive posterior density (90%ile, median and 5%ile), along with the actual isds for weeks 50, 110, 175, and 222. Weeks 50 and 175 have median isd levels, while week 110 is high, and week 222 is relatively low. It is clear from Figure 1 that the model generates isds that decrease in moneyness for the short-term options, but that the model fails to capture

some features of the data. In week 50, for example, the model fits the deepest in-the-money option, but overstates the isds of all other six options. Weeks 175 and 222 exhibit a classic smile pattern that the model does not fit. In both cases, the actual isds for options 2, 4, and 5 lie within the predictive density bands, but the deepest in-the-money and two most out-of-the-money options' isds are higher than the model would predict. In Week 110, which is characterized by relatively high isds, the entire posterior band lies below the actual isds. Recall that the at-the-money isd was lowest in week 107, so that Week 110 also follows a rapid run-up in implied volatilities in these options. The patterns in these three figures suggest that the model implies that volatility and option prices behave more smoothly than they do in practice. Turning to the intermediate term options in Figure 2 we see a slightly better fit (but we also have less extreme moneyness), with only five options. In week 110 there is a clear lack of monotonicity in the data, that is not in the model, and that transcends the predictive density bands. Here the at-the-money option (3) lies above the band, as do the deepest in- and out-of-the money options. Yet, here the fourth option lies below the band. Week 222 in an example of a data pattern that is non-monotone, but that nevertheless falls within the predictive density bands. This is the case discussed in the introduction where the model cannot fit the actual pattern of the data, but the difference between the two is small enough to attribute to parameter and state variable uncertainty.

Compare Week 50 in Figure 3 to Figure 1. We see that the model understates the isds for the longest term options (in Figure 3), while generally overstating the isds in this week for the shortest term options (in Figure 1). In week 110, where the volatilities are relatively high, the model is able to fit the longest term option's isd for the at-the-money option, and again opposite to the short term options, overstates the isds for the in- and out-of-the-money options. Thus, in general the difference between the smile patterns observed in the data and those from the model cannot be attributed to parameter and state variable uncertainty.

In general, as discussed in Das and Sundaram (1999), the smiles flatten as the option maturities lengthen. We now use the information in the full sample to examine the model's ability to fit different shaped smile patterns. We start by dividing the 249 weeks in the sample into quartiles, based on the slope of the smile on the shortest-term options (i.e., the isd of Option 1 minus the isd of Option 7). The cut-off values for these quartiles are: 3.6%,

5.3%, and 7.3%. So Group 1 comprises those 62 weeks in the sample with the smallest difference between the isd of the deepest in-the-money and most out-of-the-money shortest term options, etc. Table 6 reports the differences in isds between options 2 and 6 (short term), 8 and 12 (intermediate term), 9 and 11 (intermediate term), and 15 and 13 (longest term), within each of the quartiles defined above. The table provides the actual values and posterior properties of each of these differences, averaged over the time period within each quartile.

By comparing the differences between options 2 and 6, with those of 8 and 12, and 15 and 13, we see that the slope of the smile declines rather dramatically in the options' terms. An exception to this pattern is the first quartile for the shortest term, relative to the intermediate term options. Under the model within all four quartiles, the average smile slope for comparable moneyness pairs, is smaller the longer the options' terms. However, as noted by Das and Sundaram (1999), the magnitude of the drop in slopes is much smaller under the model than observed in the data. Consider the third quartile. Here there are 62 dates when the difference between the isd of the deepest in-the-money and deepest out-of-the-money, shortest term options falls between 5.3% and 7.3%. The average difference between option 2's isd and option 6's in these cases is 8.6%. Holding moneyness constant and increasing term by some 30 days we see that the average difference between the isds of options 8 and 12 falls to 5.9%. Under the model, the 80%ile bands for these two differences are [4.6% 4.8%] and [4.2% 4.4%]. The model understates both this difference (i.e., the slope of the smile), and the drop in the slope as term increases (the maximal model drop is 13%, compared to the actual drop of 38%). In fact, in only one of the 16 cells does the realized average smile slope lie within the 80% predictive posterior bands from the model (Group 2 – longest term: options 15 and 13). In 13 of the 16 cases, the slope of the smile under the model is smaller than that observed in the data. This inability to explain the relationship between moneyness and market price can not be attributed to parameter and state variable uncertainty, or discretization bias.

Figure 4 looks at the term structure of isds from the three at-the-money options (Options 4, 10, and 14) at the same four dates. We see that in Weeks 50, 175 and 222, where volatilities are low or average, there is an upward slope under the model–longer term options have higher

isds. This pattern also manifests in the data; although we see a steeper relationship in Week 50 in the data relative to the model. Week 110 is the case where isds are relatively high. In the data we see a downward slope with term (consistent with mean reversion), but the model has a gentle **U** shape, which as we have already seen is not capable of fitting the relatively extreme isd of the shortest term option.

6. Conclusions

This paper adds to the growing literature that jointly estimates the model parameters in both the physical and equivalent martingale measures, by developing an empirical approach that exploits all the information in a broad cross-section of option prices along with the dynamics of the underlying asset, using the exact probabilistic framework of the model. Since finance theory rests heavily on continuous time models, it is compelling to develop an empirical framework that is fully consistent with this theory–despite the fact that data are gathered in discrete time. To this end, we derive the necessary joint transition density to draw from the latent volatility process conditioned on the observed returns and option prices. While the pricing model that we estimate is generally known to be inadequate in explaining the cross-section of option prices, and the finance literature is moving on to more complicated models that include jumps in both the price and volatility of the underlying asset along with risk premia on these jumps, there remain questions along several dimensions that we answer. First, no study to date has used all the information in the exact transition densities from the stochastic volatility (Heston) model. It is possible that the documented performance of the model is influenced by discretization error. Second, in the context of efficient estimation, how much of the mispricing may be attributable to parameter and state variable uncertainty? We also explore the role of using different sets of options in estimating the model. Finally, our metric for assessing the fit of the model is the implied volatility smile. Our Bayesian approach affords an exact posterior for this important implication of the model. Using this metric we assess directly the model's ability to capture smile patterns empirically, which is an important criterion in evaluating an option pricing model.

Our results show that with fifteen and even three options, after conditioning on the model, the variance-covariance matrix of the pricing errors is not of full rank. We address

this degeneracy of the likelihood function with a prior that off-diagonal elements of the error variance-covariance matrix are 0. The Bayesian methodology allows a decomposition of the lack of perfect fit, and it allows us to cast inference in terms of the implied volatility smile. Part of the difference between the model and the data is due to (the econometrician's) lack of knowledge of the parameters and state variables. The other part is attributable to pricing errors–that under the model would be well behaved. For the deep in-the-money, short-term call, parameter and state variable uncertainty give rise to a 90%ile band in the predictive posterior density of the implied volatility of 0.56%. But the data are on average 3.6% away from the predictive posterior density mean. By contrast a slightly in-the-money, intermediate term option has average pricing errors relative to the posterior mean of 0.26%, and parameter and state variable uncertainty imply that the predictive posterior density's 90%ile band width is 0.58%–suggesting an excellent model fit.

Consider "Week 222" in Figure 2 as an example of what might have prevailed more generally. Here the pattern of Black-Scholes implied volatilities is non-monotone in moneyness, while the model suggests that these implied volatilities decline in moneyness. The pattern in the data is different from what we would expect to see under the model. Nevertheless, in this case parameter and state variable uncertainty alone create wide enough bands around this pattern to accommodate the data. By contrast, "Week 110" in Figure 1 is a case where none of the short-term options' implied volatilities fall within the model's predictive posterior bands.

By reviewing the effect that the set of options used to estimate the model has on parameter estimates, we see the role played by the volatility risk premium per unit of volatility (λ). For all of our information sets λ is negative. As the information set changes λ adapts to shift the smile vertically. For example when we add information about the smile at one term to a single option λ becomes significantly smaller, but as we add information about smiles at different terms, λ increases to its estimated level with the single option. This suggests that the model is rich enough to fit the smile, but lacks the flexibility needed to fit the term structure of smiles observed in the data.

Appendix A. Option Prices (Heston 1993).

Assuming that the return generating process is equations (1) and (2) in the text, and that the risk premium on volatility is proportional to volatility (i.e., $\lambda(\cdot) = \lambda v_t$), Heston (1993) shows that the time t value of a European call with strike price X, that expires in T years, and whose underlying asset is currently worth S (adjusted for the present value of dividends) with variance v is:

$$C(S, v, t) = SP_1 - Xe^{-rT}P_2 \tag{3}$$

where:

$$P_j(\ln(S), v, T; \ln(X)) = \frac{1}{2} + \frac{1}{\pi}\int_0^\infty Re\left[\frac{e^{-i\phi \ln(X)} f_j(\ln(S), v, T; \phi)}{i\phi}\right] d\phi \tag{4}$$

where:

$$f_j(\ln(S), v, T; \phi) = e^{C(T,\phi)+D(T;\phi)v+i\phi \ln(S)} \tag{5}$$

where:

$$C(T; \phi) = r\phi iT + \frac{a}{\sigma^2}\left\{(b_j - \rho\sigma\phi i + d)T - 2\ln\left[\frac{1 - ge^{d \cdot T}}{1 - g}\right]\right\} \tag{6}$$

and

$$D(T; \phi) = \frac{b_j - \rho\sigma\phi i + d}{\sigma^2}\left[\frac{1 - e^{dT}}{1 - ge^{dT}}\right] \tag{7}$$

and

$$g = \frac{b_j - \rho\sigma\phi i + d}{b_j - \rho\sigma\phi i - d} \tag{8}$$

and

$$d = \sqrt{(\rho\sigma\phi i - b_j)^2 - \sigma^2(2u_j\phi i - \sigma^2)} \tag{9}$$

and: $u_1 = \frac{1}{2}$, $u_2 = -\frac{1}{2}$, $a = \kappa\theta$, $b_1 = \kappa + \lambda - \rho\sigma$, $b_2 = \kappa + \lambda$

Appendix B. Derivation of the Joint Transition Density

The characteristic function for the process described by (1) and (2) in the text is:

$$\phi(k,\gamma,\tau,x_t,v_t) = E[e^{ikx_{t+\tau}+i\gamma v_{t+\tau}}|x_t,v_t] \tag{10}$$

Where v_t is the (latent) variance at time t and x_t is the logarithm of the sum of the index level and accumulated dividends since time $t=0$; (so that $r_t = x_t - x_{t-1}$).

This must satisfy the following PDE (Øksendal (1998)):

$$\frac{1}{2}v_t\frac{\partial^2\phi}{\partial x_t^2} + \rho\sigma v_t\frac{\partial^2\phi}{\partial x_t\partial v_t} + \frac{1}{2}\sigma^2 v_t\frac{\partial^2\phi}{\partial v_t^2} + (\mu - \frac{1}{2}v_t)\frac{\partial\phi}{\partial x_t} + \kappa(\theta - v_t)\frac{\partial\phi}{\partial v_t} - \frac{\partial\phi}{\partial\tau} = 0 \tag{11}$$

and,

$$\phi(k,\gamma,0,x_T,v_T) = e^{ikx_T+i\gamma v_T} \tag{12}$$

Given this affine structure, we look for a solution of the form:

$$\phi = e^{ikx_t+A(k,\tau,\gamma)v_t+B(k,\tau,\gamma)} \tag{13}$$

with boundary conditions:

$$A(k,0,\gamma) = i\gamma \tag{14}$$

$$B(k,0,\gamma) = 0 \tag{15}$$

In this case, the PDE becomes:

$$\frac{1}{2}v_t(ik)^2 + \rho\sigma v_t ikA + \frac{1}{2}\sigma^2 v_t A^2 + (\mu - \frac{1}{2}v_t)ik + \kappa(\theta - v_t)A - (\dot{A}v_t + \dot{B}) = 0 \tag{16}$$

where, $\dot{A} = \frac{dA}{dt}$.

Let:

$$\begin{aligned}
a &= \frac{\sigma^2}{2} \\
b &= ik\rho\sigma - \kappa \\
c &= -\left(\frac{k^2}{2} + i\frac{k}{2}\right) \\
d &= \sqrt{(\kappa - ik\rho\sigma)^2 + \sigma^2(k^2 + ik)} \\
\zeta &= ik\mu \\
f &= \kappa\theta \\
A_1 &= \frac{1}{2a}\left[-b + \sqrt{b^2 - 4ac}\right] = \frac{1}{2a}[-b + d] \\
A_2 &= \frac{1}{2a}\left[-b - \sqrt{b^2 - 4ac}\right] = \frac{1}{2a}[-b - d]
\end{aligned}$$

Functions A and B then satisfy the following Riccatti equations:

$$\dot{A} = aA^2 + bA + c \tag{17}$$

$$\dot{B} = \zeta + fA \tag{18}$$

The solution for A is given by

$$A = \frac{A_2 e^{d(\tau+\xi)} - A_1}{e^{d(\tau+\xi)} - 1} \tag{19}$$

where ξ is the constant of integration.

Let $g(\gamma) = \frac{i\gamma - A_1}{i\gamma - A_2}$, then:

$$A = \frac{A_2 e^{d\tau} g(\gamma) - A_1}{e^{d\tau} g(\gamma) - 1} \tag{20}$$

$$\dot{B} = \zeta + f\frac{A_2 g(\gamma) e^{d\tau} - A_1}{g(\gamma) e^{d\tau} - 1} = \zeta + f\left[A_2 + \frac{A_2 - A_1}{g(\gamma) e^{d\tau} - 1}\right] \tag{21}$$

$$B = \zeta + fA_2 - \frac{df/a}{g(\gamma) e^{d\tau} - 1}. \tag{22}$$

Implying:

$$B = \zeta + (\zeta + fA_2)\tau - \frac{f}{a} ln(e^{-d\tau} - g(\gamma)) \tag{23}$$

$$= (\zeta + fA_2)\tau - \frac{f}{a} ln\left[\frac{e^{-d\tau} - g(\gamma)}{1 - g(\gamma)}\right]. \tag{24}$$

And the boundary conditions:

$$A(\tau = 0) = i\gamma \tag{25}$$

$$B(\tau = 0) = 0 \tag{26}$$

Thus,

$$\phi(k, \gamma, \tau, x, v) = e^{ikx + A(k,\tau,\gamma)v + B(k,\tau,\gamma)} \tag{27}$$

So the joint conditional density transition density is:

$$G(x_{t+\tau}, v_{t+\tau}|x_t, v_t) = \frac{1}{4\pi^2} \iint e^{-ikx_{t+\tau} - i\gamma v_{t+\tau}} \phi(k, \gamma, \tau, x_t, v_t)\, dk\, d\gamma \tag{28}$$

Our objective is to use the special form of the diffusion to obtain a "closed-form" solution to one of these integrals. To this end we re-write the double integral as follows:

$$\int \left\{ \frac{1}{2\pi} e^{ik(x_{t+\tau} - x_t)} \int \left\{ \frac{1}{2\pi} e^{-i\gamma v_{t+\tau} + A + Bv_t} \right\} d\gamma \right\} dk \tag{29}$$

and collect the terms containing γ:

$$e^{-i\gamma v_{t+\tau}} e^{\frac{A_2 g(\gamma) e^{d\tau} - A_1}{e^{d\tau}\gamma - 1} v_t + (\zeta + fA_2)\tau - \frac{f}{a} ln\left[\frac{e^{-d\tau} - g(\gamma)}{1 - g(\gamma)}\right]} \tag{30}$$

$$= e^{-i\gamma v_{t+\tau}} e^{\frac{A_2 g(\gamma) e^{d\tau} - A_1}{e^{d\tau}\gamma - 1} v_t + (ik\mu + \kappa\theta A_2)\tau - (\nu+1) ln[\frac{e^{-d\tau} - g(\gamma)}{1 - g(\gamma)}]} \tag{31}$$

$$\propto e^{-i\gamma v_{t+\tau}} \left(\frac{e^{-d\tau} - g(\gamma)}{1 - g(\gamma)} \right)^{\nu+1} e^{\frac{A_2 g(\gamma) e^{d\tau} - A_1}{e^{d\tau} g(\gamma) - 1} v_t}. \tag{32}$$

Where, $\nu = \frac{2\kappa\theta}{\sigma^2} - 1$.

Further simplifying:

$$\frac{1}{g(\gamma)e^{d\tau} - 1} = \frac{1}{\frac{i\gamma - A_1}{i\gamma - A_2} e^{d\tau} - 1} = \frac{i\gamma - A_2}{i\gamma(e^{d\tau} - 1) - A_1 e^{d\tau} + A_2} \tag{33}$$

And with some tedious algebra, we find:

$$\frac{e^{-d\tau} - g(\gamma)}{1 - g(\gamma)} = \frac{A_1 - A_2 e^{-d\tau}}{d/a} + \frac{e^{-d\tau} - 1}{d/a} i\gamma \tag{34}$$

and:

$$\frac{1}{g(\gamma)e^{d\tau} - 1} = \frac{1}{e^{d\tau} - 1} + \frac{(\psi + \frac{d}{a})\frac{1}{e^{d\tau} - 1}}{i\gamma - A_1 - \psi} \tag{35}$$

Where $\psi = \frac{d}{a(e^{d\tau}-1)}$.

So the integral over terms involving γ becomes:

$$e^{(fA_2+\zeta)\tau-\frac{f}{a}ln(\frac{e^{-d\tau}}{\psi})+[A_2-\psi]v_t}\int\left\{\frac{1}{2\pi}e^{-i\gamma v_{t+\tau}-\frac{f}{a}ln[\psi e^{d\tau}+A_2-i\gamma]-\frac{\varphi v_t}{i\gamma-\psi-A_1}}\right\}d\gamma. \tag{36}$$

where $\varphi = \left(\frac{d}{a}+\psi\right)\psi = \psi^2 e^{d\tau}$.

Using a change of variable and defining $\lambda = Re[\psi + A_1]$, we can rewrite the integral in the last expression as:

$$\int_{\lambda-i\infty}^{\lambda+i\infty}\left\{\frac{1}{2\pi i}e^{i\tilde{\gamma}v_{t+\tau}}e^{-(\psi+A_1)v_{t+\tau}}e^{-\frac{f}{a}lni\tilde{\gamma}+\frac{\varphi v_t}{i\tilde{\gamma}}}\right\}d\tilde{\gamma} \tag{37}$$

This exists analytically:

$$e^{-(\psi+A_1)v_{t+\tau}}\left(\frac{v_{t+\tau}}{\varphi v_t}\right)^{\frac{\nu}{2}}I_\nu\left[2\sqrt{\varphi v_t v_{t+\tau}}\right]. \tag{38}$$

where $I_\nu[\,\cdot\,]$ is the modified Bessel function of order ν.

Thus the joint density may be written as follows:

$$G(x_{t+\tau}, v_{t+\tau}|x_t, v_t) = \frac{1}{2\pi}\left(\frac{v_{t+\tau}}{v_t}\right)^{\nu/2}\int_{-\infty}^{+\infty}dk\ \psi\ e^{f(k)}I_\nu\left[2\sqrt{\varphi v_t v_{t+\tau}}\right] \tag{39}$$

where

$$\begin{aligned} f(k) &= -ik(x_{t+\tau}-x_t-\mu\tau)+\frac{\nu+1}{2}(\kappa-ik\rho\sigma)\tau+ \\ &\quad \frac{\tau d}{2}+(A_2-\psi)v_t-(A_1+\psi)v_{t+\tau} \end{aligned}$$

Appendix C. Full Conditional Densities

From Appendix B, we see that:

$$f(v_t \mid x_{t+\tau}, x_t, v_{t+\tau}, v_{t-\tau}, \boldsymbol{\theta}) \propto G(x_{t+\tau}, v_{t+\tau} \mid x_t, v_t, \boldsymbol{\theta})\cdot G(x_t, v_t \mid x_{t-\tau}, v_{t-\tau}, \boldsymbol{\theta})\cdot \mathcal{L}(\mathcal{C}_t \mid v_t, \boldsymbol{\theta}, \boldsymbol{\Upsilon}) \tag{40}$$

$G(\ )$ is the function of Equation 33 in Appendix B. $\mathcal{C}_t$ is the 15-vector of call prices observed at time t. $\boldsymbol{\Upsilon}$ is the $(15\ \times\ 4)$ matrix containing the underlying S & P 500 value,

time-to-maturity, interest rate, and dividends corresponding to each element in $\mathcal{C}_t$. $\boldsymbol{\theta}$ is the vector of model parameters and state variables: $\{\theta, \kappa, \sigma, \rho, \lambda, \mu, \{v\}\}$.

$\mathcal{L}$ is the portion of the likelihood from the option pricing model:

$$\mathcal{L}_t(\mathcal{C}_t \mid v_t, \boldsymbol{\theta}, \boldsymbol{\Upsilon}) \propto |\Sigma|^{-.5} e^{-.5(\mathcal{C}-C_t)'\Sigma^{-1}((\mathcal{C}-C_t)} \tag{41}$$

C_t is the model value for the call price, as a function of the model parameters and v_t, as shown in Appendix B.

As described in the text, we also integrate over v_0 and v_{T+1} by drawing them as follows:

$$f(v_0|v_1, \cdot) \propto G(x_1, v_1|x_0, v_0) \tag{42}$$

where x_1, v_1, and x_0 are fixed.

And:

$$f(v_{T+1}|v_T, \cdot) \propto G(x_{T+1}, v_{T+1}|x_T, v_T) \tag{43}$$

where x_T, v_T, and x_{T+1} are fixed.

Similarly, the full conditional density for θ is proportional to:

$$f(\theta|\cdot) \propto \prod_{t=1}^{T+1} G(x_t, v_t|x_{t-1}, v_{t-1})\mathcal{L}_t \tag{44}$$

where all parameters and state variables other than θ are considered fixed, and this expression is a kernel for θ.

The full conditional densities for κ, ρ, and σ are identical in structure.

References

Aït-Sahalia, Yacine and Robert Kimmel, 2007, Maximum likelihood estimation of stochastic volatility models, *Journal of Financial Economics* 83, 413–452.

Aït-Sahalia, Yacine, Yubu Wang and Francis Yared, 2001, Do option markets correctly price the probabilities of movement of the underlying asset?, *Journal of Econometrics* 102, 67–110.

Bates, David S., 1996, Jumps and stochastic volatility: Exchange rate processes implicit in Deutsche Mark options, *Review of Financial Studies* 9, 69–107.

Black, Fisher and Myron S. Scholes, 1973, The pricing of options and corporate liabilities, *Journal of Political Economy* 81, 637–654.

Broadie, Mark, Mikhail Chernov, Michael Johannes, 2005, Model specification and risk premia: Evidence from futures options, Working Paper, Columbia University.

Broadie, Mark and Özgür Kaya, 2004, Exact simulation of stochastic volatility and other affine jump diffusion processes, Working Paper, Columbia University.

Chen, Ren-Raw and Louis Scott, 1993, Maximum Likelihood Estimation for a Multifactor Equilibrium Model of the Term Structure of Interest Rates, *Journal of Fixed Income* 3, 14–31.

Chernov, Mikhail and Eric Ghysels, 2000, A study towards a unified approach to the joint estimation of objective and risk neutral measures for the purpose of options valuation, *Journal of Financial Economics* 56, 407–458.

Chib, Siddhartha and Edward Greenberg, 1994, Bayes Inference for Regression Models with ARMA(p, q) Errors, *Journal of Econometrics*, 64, 183-206.

Chib, Siddhartha and Edward Greenburg, 1995, Understanding the Metropolis-Hastings algorithm, *The American Statistician* 49, 327–335.

Chib, Siddhartha, Michael Pitt, and Neil Shephard, 2004, Likelihood based inference for diffusion driven models, Working Paper.

Christoffersen, Peter and Kris Jacobs, 2004, The importance of the loss function in option valuation, *Journal of Financial Economics* 72, 291–318.

Cox, John, Jonathan Ingersoll, and Stephen Ross, 1985, "A Theory of the Term Structure of Interest Rates," *Econometrica,* 53, 385–407.

Das, Sanjiv Ranjan and Rangarajan K. Sundaram, 1999, Of smiles and smirks: A term structure perspective, *Journal of Financial and Quantitative Analysis* 34, 211–239.

Elerian, Ola, Siddhartha Chib, and Neil Shephard, 2001, Likelihood inference for discretely observed nonlinear diffusions, *Econometrica* 69, 959-994.

Eraker, Bjorn, 2004, Do stock prices and volatility jump? Reconciling evidence from spot and option prices, *Journal of Finance* 59, 1367–1403.

Eraker, Bjorn, Michael Johannes, and Nicholas Polson, 2003, The impact of jumps in equity index volatility and returns, *Journal of Finance* 58, 1269-1300.

Feller, William, 1951, Two singular diffusion problems, *Annals of Mathematics* 54, 173–182.

Gallant, A. Ronald, and George E. Tauchen, 1996, Which moments to match?, *Econometric Theory* 12, 657-681.

Gallant, A. Ronald, and George E. Tauchen, 1998, Reprojecting partially observed systems with application to interest rate diffusions, *Journal of the American Statistical Association* 93, 10-24.

Garcia, René, Eric Ghysels and Eric Renault, 2003, The Econometrics of Option Pricing, forthcoming in Y. Ait-Sahalia and L.P. Hansen (eds.), *The Handbook of Financial Econometrics.*

Geman, Hélyette and Marc Yor, 1993, Bessel processes, Asian options and perpetuities, *Mathematical Finance* 3, 349–375.

Heston, Steve, 1993, A closed-form solution for options with stochastic volatility with applications to bond and currency options, *Review of Financial Studies* 6, 327–343.

Hull, John and Alan White, 1987, The pricing of options on assets with stochastic volatilities, *Journal of Finance* 42, 281–300.

Jacquier, Eric, and Robert Jarrow, 2000, Bayesian analysis of contingent claim model error, *Journal of Econometrics* 94, 145–180.

Jacquier, Eric, Nicholas Polson, and Peter Rossi, 1994, Bayesian analysis of stochastic volatility models, *Journal of Business and Economic Statistics* 12, 371–418.

Jacquier, Eric, Nicholas Polson, and Peter Rossi, 2004, Bayesian analysis of stochastic volatility models with fat tails and correlated errors, *Journal of Econometrics* 122, 185–212.

Johannes, Michael S., and Nicholas Polson, 2003,MCMC methods for continuous-time financial econometrics. Available at SSRN: http://ssrn.com/abstract=480461 or DOI: 10.2139/ssrn.480461

Jones, Christopher S., 2003, The dynamics of stochastic volatility: Evidence from underlying and options markets, *Journal of Econometrics* 116, 181–224.

Judd, K. L., 1998, *Numerical Methods in Economics*, MIT Press.

Kloeden, Peter, and Eckhard Platen, 1999, *Numerical Solution of Stochastic Differential Equations*, Springer-Verlag, Berlin.

Lamoureux, Christopher G. and William Lastrapes, 1993, Forecasting stock return variance: Toward an understanding of stochastic implied volatilities, *Review of Financial Studies* 6, 293–326.

Lamoureux, Christopher G. and H. Douglas Witte, 2002, Empirical analysis of the yield curve: The information in the data viewed through the window of Cox, Ingersoll, and Ross, *Journal of Finance* 57, 1479–1520.

Muirhead, Robb J., 1982, *Aspects of Multivariate Statistical Theory*, John Wiley & Sons.

Øksendal, Bernt, 1998, *Stochastic differential equations: An introduction with applications*, Springer.

Pan, Jun, 2002, The jump-risk premia implicit in options: Evidence from an integrated time-series study, *Journal of Financial Economics* 63, 3–50.

Pearson, Neil and Tong-Sheng Sun, 1994, Exploiting the conditional density in estimating the term structure: an application to the Cox, Ingersoll, and Ross model, *Journal of Finance* 49, 1279-1304.

Revuz, Daniel and Marc Yor, 1999, *Continuous Martingales and Brownian Motion (3rd ed.)*, Springer-Verlag, Berlin.

Robert, Christian P. and George Casella, 2004, *Monte Carlo Statistical Methods (2nd ed.)*, Springer, New York.

Schotman, Peter, 1996, A Bayesian approach to the empirical evaluation of bond options, *Journal of Econometrics* 75, 183–215.

Sims, Christopher A., 2003, Comment (On Iterative and recursive estimation in structural non-adaptive models), *Journal of Business and Economic Statistics* 21, 500-503.

Singleton, Kenneth J., 2001, Estimation of affine asset pricing models using the empirical characteristic function, *Journal of Econometrics* 102, 111–141.

Singleton, Kenneth J., 2006, *Empirical Dynamic Asset Pricing*, Princeton University Press, Princeton.

Table 1
Properties of the Joint Density if v_{t+1}: Continuous and Discrete Cases

Case			1%ile	5%ile	25%ile	Median	75%ile	95%ile	99%ile	Maximum	$\tilde{v}_{t+1}$
P_{t+1}	v_t	Method								CDF (kernel)	at max
450	0.01	GF	.0274	.0305	.0363	.0408	.0455	.0516	.0540	0.4E-3	.0403
450	0.01	ED	.0206	.0239	.0286	.0318	.0350	.0397	.0429	0.3E-11	.0318
450	0.015	GF	.0287	.0325	.0389	.0435	.0481	.0530	.0544	.004	.0437
450	0.015	ED	.0229	.0269	.0326	.0366	.0406	.0463	.0503	0.4E-7	.0366
450	0.025	GF	.0335	.0385	.0464	.0523	.0586	.0684	.0757	0.06	.0514
450	0.025	ED	.0286	.0338	.0411	.0462	.0513	.0587	.0639	0.009	.0462
450	0.029	GF	.0358	.0410	.0493	.0555	.0622	.0724	.0801	0.1	.0546
450	0.029	ED	.0311	.0366	.0446	.0501	.0556	.0635	.0691	0.003	.0501
450	0.033	GF	.0382	.0437	.0524	.0588	.0657	.0764	.0844	0.2	.0579
450	0.033	ED	.0336	.0396	.0480	.0539	.0598	.0682	.0742	0.2	.0539
450	0.43	GF	.0445	.0507	.0602	.0673	.0749	.0866	.0953	0.6	.0664
450	0.43	ED	.0404	.0472	.0568	.0635	.0702	.0799	.0867	1.2	.0635
450	0.48	GF	.0478	.0543	.0642	.0717	.0795	.0917	.1007	0.9	.0707
450	0.48	ED	.0439	.0511	.0612	.0683	.0754	.0856	.0928	2.4	.0683
475	0.01	GF	.0125	.0156	.0205	.0244	.0286	.0355	.0408	4.6	.0236
475	0.01	ED	.0099	.0131	.0178	.0210	.0242	.0289	.0321	1.0	.0210
475	0.015	GF	.0149	.0184	.0240	.0284	.0332	.0408	.0467	92	.0275
475	0.015	ED	.0121	.0161	.0218	.0258	.0298	.0355	.0395	10	.0258
475	0.025	GF	.0204	.0247	.0316	.0369	.0426	.0516	.0585	25	.0359
475	0.025	ED	.0178	.0229	.0303	.0354	.0405	.0479	.0531	64	.0354
475	0.029	GF	.0229	.0275	.0348	.0404	.0465	.0560	.0632	30	.0394
475	0.029	ED	.0202	.0258	.0338	.0393	.0448	.0527	.0583	94	.0393
475	0.033	GF	.0254	.0303	.0380	.0440	.0504	.0603	.0679	34	.0430
475	0.033	ED	.0228	.0288	.0372	.0431	.0490	.0574	.0634	125	.0431
475	0.043	GF	.0320	.0376	.0464	.0530	.0601	.0712	.0796	42	.0520
475	0.043	ED	.0296	.0364	.0460	.0527	.0594	.0655	.0759	203	.0527
475	0.048	GF	.0355	.0414	.0506	.0576	.0650	.0766	.0853	44	.0566
475	0.048	ED	.0331	.0402	.0504	.0575	.0646	.0748	.0820	240	.0575
500	0.01	GF	.0023	.0040	.0074	.0102	.0136	.0194	.0241	1000	.0091
500	0.01	ED	-.0004	.0028	.0075	.0107	.0140	.0186	.0219	1000	.0107
500	0.015	GF	.0047	.0071	.0114	.0150	.0190	.0258	.0311	674	.0138
500	0.015	ED	.0019	.0059	.0116	.0155	.0195	.0252	.0292	1000	.0155
500	0.025	GF	.0104	.0139	.0197	.0244	.0296	.0379	.0444	408	.0233
500	0.025	ED	.0075	.0127	.0200	.0252	.0303	.0376	.0428	1000	.0252
500	0.029	GF	.0128	.0167	.0231	.0282	.0337	.0426	.0494	352	.0270
500	0.029	ED	.0100	.0156	.0235	.0290	.0345	.0424	.0480	1000	.0290
500	0.033	GF	.0154	.0196	.0265	.0319	.0378	.0472	.0544	310	.0308
500	0.033	ED	.0126	.0185	.0270	.0328	.0387	.0472	.0531	1000	.0328
500	0.043	GF	.0221	.0271	.0352	.0414	.0481	.0586	.0666	238	.0402
500	0.043	ED	.0193	.0261	.0358	.0425	.0492	.0588	.0656	1000	.0425
500	0.048	GF	.0256	.0310	.0395	.0461	.0531	.0642	.0726	214	.0450
500	0.048	ED	.0228	.0300	.0402	.0473	.0544	.0646	.0717	1000	.0473

Table 1 (Cont'd.)
Properties of the Joint Density of v_{t+1}: Continuous and Discrete Cases

Case			1%ile	5%ile	25%ile	Median	75%ile	95%ile	9 9%ile	Maximum	$\tilde{v}_{t+1}$
P_{t+1}	v_t	Method								CDF (kernel)	at max
525	0.01	GF	.0018	.0032	.0061	.0086	.0117	.0171	.0216	1	.0075
525	0.01	ED	-.0102	-.0069	-.0023	.0010	.0042	.0089	.0121	2	.0010
525	0.015	GF	.0025	.0043	.0079	.0110	.0147	.0184	.0260	8	.0097
525	0.015	ED	-.0079	-.0039	.0018	.0058	.0097	.0154	.0194	16	.0058
525	0.025	GF	.0059	.0087	.0138	.0180	.0228	.0306	.0368	35	.0168
525	0.025	ED	-.0022	.0029	.0103	.0154	.0205	.0279	.0330	84	.0154
525	0.029	GF	.0078	.0110	.0167	.0213	.0264	.0348	.0414	44	.0200
525	0.029	ED	.0002	.0058	.0137	.0192	.0248	.0327	.0382	118	.0192
525	0.033	GF	.0099	.0135	.0197	.0247	.0302	.0391	.0460	51	.0235
525	0.033	ED	.0028	.0088	.0172	.0231	.0290	.0374	.0434	152	.0231
525	0.043	GF	.0158	.0203	.0277	.0335	.0399	.0500	.0577	63	.0323
525	0.043	ED	.0096	.0163	.0260	.0327	.0394	.0491	.0558	236	.0327
525	0.048	GF	.0190	.0239	.0319	.0380	.0448	.0554	.0635	65	.0369
525	0.048	ED	.0131	.0202	.0304	.0375	.0446	.0548	.0620	274	.0375
550	0.025	GF	.0072	.0100	.0151	.0192	.0239	.0315	.0375	0.005	.0185
550	0.025	ED	-.0116	-.0064	.0010	.0061	.0112	.0186	.0237	0.08	.0061
550	0.029	GF	.0082	.0114	.0168	.0213	.0263	.0344	.0407	0.02	.0205
550	0.029	ED	-.0091	-.0035	.0044	.0099	.0154	.0234	.0289	0.3	.0099
550	0.033	GF	.0095	.0129	.0188	.0236	.0289	.0376	.0443	0.07	.0225
550	0.033	ED	-.0065	-.0005	.0079	.0138	.0197	.0281	.0341	0.8	.0138
550	0.043	GF	.0137	.0179	.0249	.0305	.0366	.0464	.0539	0.5	.0291
550	0.043	ED	.0003	.0070	.0167	.0234	.0301	.0398	.0465	4	.0234
550	0.048	GF	.0162	.0208	.0284	.0343	.0408	.0511	.0590	0.9	.0329
550	0.048	ED	.0038	.0109	.0211	.0282	.0353	.0455	.0527	7	.0010
550	0.055	GF	.0201	.0252	.0335	.0399	.0469	.0579	.0663	2	.0387
550	0.055	ED	.0088	.0164	.0274	.0349	.0425	.0534	.0611	13	.0349

This table reports properties of the cumulative density function for $f(x_{t+1}, \tilde{v}_{t+1}|x_t, v_t)$, where x_t is $\ln(P_t)$, the level of the index at t; and v_t is the variance at t. For the continuous case, we use the joint transition density derived in Appendix B to construct this cdf of v_{t+1} treating everything else as fixed. Model parameters and x_0 are fixed as: $\theta = 0.029$; $\kappa = 2.9$; $\sigma_v = 0.4$; $\rho = -0.5$; $\mu = 0.08; \delta t = \frac{7}{365}$; $x_t = 500$.

Table 2
Properties of the Options Data

Option	Moneyness (S/X)				Term (in Years)				Implied Standard Deviation			
Number	Mean	Std Dev	Minimum	Maximum	Mean	Std Dev	Minimum	Maximum	Mean	Std Dev	Minimum	Maximum
1	1.04299	0.02815	1.02452	1.46878	0.05968	0.02435	0.02466	0.10685	0.19265	0.05169	0.00101	0.36463
2	1.02682	0.00598	1.01331	1.04489	0.05968	0.02435	0.02466	0.10685	0.17857	0.04528	0.10069	0.34828
3	1.01276	0.00469	1.00318	1.02668	0.05968	0.02435	0.02466	0.10685	0.16319	0.04126	0.09757	0.31661
4	0.99903	0.00406	0.99005	1.00886	0.05968	0.02435	0.02466	0.10685	0.15138	0.03869	0.09120	0.29377
5	0.98573	0.00413	0.97465	0.99304	0.05968	0.02435	0.02466	0.10685	0.14369	0.03598	0.08729	0.28120
6	0.97274	0.00493	0.95813	0.98237	0.05968	0.02435	0.02466	0.10685	0.13843	0.03387	0.08349	0.25502
7	0.95983	0.00685	0.92091	0.97336	0.05968	0.02435	0.02466	0.10685	0.13781	0.03273	0.08072	0.26025
8	1.02330	0.00564	1.00902	1.03844	0.15449	0.05465	0.10137	0.35342	0.17469	0.04171	0.11142	0.32819
9	1.00926	0.00463	0.99630	1.02054	0.15449	0.05465	0.10137	0.35342	0.16628	0.04032	0.10496	0.31411
10	0.99560	0.00432	0.98168	1.00331	0.15449	0.05465	0.10137	0.35342	0.15816	0.03874	0.09854	0.29848
11	0.98232	0.00478	0.96658	0.99131	0.15449	0.05465	0.10137	0.35342	0.15100	0.03686	0.09515	0.28301
12	0.96938	0.00563	0.95196	0.98089	0.15449	0.05465	0.10137	0.35342	0.14520	0.03488	0.09018	0.25686
13	1.00619	0.00968	0.98836	1.08238	0.30653	0.11047	0.17808	0.60274	0.16931	0.03686	0.11343	0.27981
14	0.99044	0.00672	0.95093	1.00059	0.30653	0.11047	0.17808	0.60274	0.16277	0.03614	0.10597	0.27249
15	0.97550	0.01381	0.88335	0.98841	0.30653	0.11047	0.17808	0.60274	0.15670	0.03523	0.10052	0.26617

The data are weekly from January 3, 1989 through October 6, 1993, European Call Options on the S&P 500 Index.
The level of the index at the beginning of the period was approximately 276.77 and it was 460.26 at the end of the period.

Table 3
Error Correlation Matrix
The Case of the Diffuse Prior

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	**0.981**	0.969	0.941	0.926	0.894	0.862	0.629	0.873	0.863	0.835	0.805	0.790	0.594	0.531	0.463
2	0.969	**0.969**	0.964	0.940	0.903	0.862	0.636	0.885	0.877	0.852	0.823	0.802	0.609	0.545	0.476
3	0.941	0.964	**0.829**	0.963	0.923	0.862	0.638	0.885	0.879	0.863	0.834	0.810	0.623	0.567	0.502
4	0.926	0.940	0.963	**0.653**	0.950	0.881	0.680	0.862	0.856	0.844	0.816	0.794	0.594	0.536	0.464
5	0.894	0.903	0.923	0.950	**0.479**	0.932	0.760	0.803	0.793	0.785	0.760	0.743	0.528	0.472	0.396
6	0.862	0.862	0.862	0.881	0.932	**0.303**	0.837	0.721	0.701	0.682	0.644	0.628	0.382	0.330	0.236
7	0.629	0.636	0.638	0.680	0.760	0.837	**0.243**	0.542	0.532	0.522	0.503	0.504	0.250	0.206	0.139
8	0.873	0.885	0.885	0.862	0.803	0.721	0.542	**1.364**	0.984	0.968	0.948	0.925	0.796	0.749	0.691
9	0.863	0.877	0.879	0.856	0.793	0.701	0.532	0.984	**1.251**	0.978	0.962	0.938	0.802	0.758	0.701
10	0.835	0.852	0.863	0.844	0.785	0.682	0.522	0.968	0.978	**1.077**	0.979	0.956	0.798	0.759	0.710
11	0.805	0.823	0.834	0.816	0.760	0.644	0.503	0.948	0.962	0.979	**0.892**	0.968	0.798	0.759	0.718
12	0.790	0.802	0.810	0.794	0.743	0.628	0.504	0.925	0.938	0.956	0.968	**0.681**	0.753	0.718	0.672
13	0.594	0.609	0.623	0.594	0.528	0.382	0.250	0.796	0.802	0.798	0.798	0.753	**1.446**	0.976	0.953
14	0.531	0.545	0.567	0.536	0.472	0.330	0.206	0.749	0.758	0.759	0.759	0.718	0.976	**1.345**	0.974
15	0.463	0.476	0.502	0.464	0.396	0.236	0.139	0.691	0.701	0.710	0.718	0.672	0.953	0.974	**1.191**

Notes: This table reports posterior mean error standard deviations along the diagonal (in bold), and correlations on the off-diagonal.
This is a (steady-state) draw from the error variance-covariance matrix when a diffuse prior is used.
As explained in the text, this matrix is technically of full rank, but virtually singular. In this case,
all information would come from option prices, and the model would be unidentified. for this reason, the analysis
in the paper is conducted using a prior that Σ is diagonal.

As described in the text, Options 1 - 7 are shortest term to maturity.
Options 8 - 12 are intermediate term to maturity.
Options 13 - 15 are long term to maturity.
Within term options start from deepest-in-the-money.
Thus, Option 7 is the deepest-out-of-the-money, short-term option, and Option 8 is the deepest in-the-money, intermediate option.

Table 4
Parameter Posteriors
The Case of the Diagonal Prior on Σ

	5%ile	10%ile	25%ile	Median	Mean	75%ile	90%ile	95%ile	std dev
θ	0.02066	0.02140	0.02264	0.02406	0.02560	0.02748	0.03285	0.03570	0.00435
	0.01897	0.01913	0.02036	0.02116	0.02142	0.02288	0.02363	0.02418	0.00170
	0.02162	0.02242	0.02396	0.02623	0.02720	0.02907	0.03345	0.03800	0.00479
	0.02207	0.02325	0.02546	0.02873	0.03026	0.03302	0.03966	0.04442	0.00706
κ	2.10410	2.29043	2.68075	3.04738	2.96295	3.28177	3.45603	3.54074	0.42944
	7.66094	7.85339	8.26764	8.83700	8.85198	9.49245	9.97753	10.22713	0.78730
	2.75098	2.99221	3.44661	3.95766	3.97479	4.50110	4.99666	5.27599	0.75976
	1.69677	1.94194	2.45173	3.01576	3.00426	3.54569	4.01976	4.29683	0.78589
σ	0.37432	0.37625	0.37974	0.38373	0.38372	0.38772	0.39108	0.39301	0.00574
	0.58292	0.58828	0.59668	0.60623	0.60526	0.61378	0.62092	0.62541	0.01324
	0.23993	0.24792	0.26187	0.27819	0.27939	0.29632	0.31246	0.32207	0.02493
	0.20879	0.21490	0.22609	0.23907	0.23996	0.25292	0.26582	0.27376	0.01973
ρ	-.83828	-.83542	-.83018	-.82395	-.82370	-.81751	-.81149	-.80807	0.00923
	-.73824	-.73280	-.72465	-.71603	-.71621	-.70748	-.69995	-.69557	0.01281
	-.70793	-.69090	-.66214	-.62873	-.62637	-.59234	-.55837	-.53752	0.05173
	-.77234	-.75925	-.73570	-.70736	-.70504	-.67689	-.64796	-.62936	0.04362
μ	0.01930	0.03533	0.06049	0.08856	0.08823	0.11661	0.14096	0.15538	0.04135
	0.00829	0.02424	0.04950	0.08106	0.08199	0.11294	0.14178	0.15928	0.04602
	-.06270	-.04256	-.00929	0.02706	0.02758	0.06408	0.09843	0.11875	0.05477
	-.08268	-.06464	-.03101	0.00542	0.00607	0.04305	0.07724	0.09795	0.05514
λ	-2.15833	-2.05146	-1.88028	-1.67915	-1.57039	-1.29557	-0.86447	-0.67738	0.43126
	-3.92075	-3.80632	-3.53460	-3.07968	-3.12956	-2.70709	-2.50710	-2.42943	0.48441
	-2.71110	-2.54418	-2.20669	-1.77546	-1.74339	-1.36207	-0.86652	-0.48342	0.65195
	-2.55340	-2.32622	-1.96081	-1.52020	-1.49999	-1.07949	-0.62361	-0.34791	0.66223
$\frac{2\kappa\theta}{\sigma^2}$	1.00030	1.00061	1.00163	1.00383	1.00540	1.00756	1.01233	1.01577	0.00519
	1.00150	1.00302	1.00886	1.02131	1.02853	1.04142	1.06449	1.08189	0.02577
	2.23761	2.33026	2.49210	2.69817	2.71979	2.92557	3.13757	3.27122	0.31880
	2.41452	2.54420	2.75706	3.00126	3.01338	3.25619	3.49995	3.65662	0.37801
ss var	0.02349	0.02440	0.02598	0.02788	0.03001	0.03246	0.03986	0.04396	0.00597
	0.01931	0.01966	0.02062	0.02188	0.02206	0.02357	0.02462	0.02517	0.00186
	0.00990	0.01046	0.01155	0.01301	0.01380	0.01510	0.01793	0.02128	0.00342
	0.00979	0.01054	0.01202	0.01430	0.01601	0.01791	0.02365	0.02824	0.00621
θ^Q	0.05074	0.05127	0.05215	0.05316	0.05324	0.05427	0.05532	0.05594	0.00158
	0.02976	0.03038	0.03140	0.03293	0.03308	0.03445	0.03600	0.03694	0.00226
	0.04012	0.04148	0.04402	0.04736	0.04879	0.05168	0.05725	0.06209	0.00965
	0.04441	0.04661	0.05102	0.05779	0.06505	0.06847	0.08565	0.10376	0.34734
κ^Q	1.27319	1.29857	1.34263	1.39168	1.39256	1.44171	1.48745	1.51391	0.07324
	4.98179	5.15569	5.39915	5.72495	5.72241	6.08157	6.29777	6.40671	0.45523
	1.39072	1.57203	1.88810	2.23327	2.23140	2.57370	2.87411	3.06235	0.51449
	0.61834	0.80745	1.13104	1.50297	1.50427	1.86909	2.20253	2.39421	0.54351

Table 4
Parameter Posteriors
The Case of the Diagonal Prior on Σ (Cont'd.)

Notes: This table reports aspects of the posterior densities of the model parameters and functions of interest. The model is estimated using weekly data from 1989 through October 1993, on four different data sets and the underlying asset (S&P 500 index).
The four information sets are: 1) all 15 options; 2) Options 8, 10, and 12; 3) Option 10; and 4) Option 10, with $\Sigma_{(1,1)}$ fixed at 0.2.

The model we consider for the evolution of stock returns is:

$$dS = \mu S dt + \sqrt{v} S dz$$
$$dv = \kappa(\theta - v)dt + \sigma\sqrt{v} d\omega$$

The model for option prices is: $P(t, X, T) = C(t, X, T, \cdot) + \epsilon$.
$\epsilon \sim N(0, \Sigma)$.

As explained in the text, identification is achieved by placing a diagonal prior on Σ.

Table 5
The relationship between the actual ISDs and the Predictive Densities

Option	Root-Mean Sq. Error - Mean	Root-Mean Sq. Error - Median	Average 90%ile band	# Sample Below band	# Sample Above band
1	3.599	3.599	0.568	39	193
	6.547	7.137	0.773	54	164
	7.681	7.678	1.786	8	208
	6.569	5.790	2.162	8	206
2	2.171	2.171	0.582	48	166
	2.071	2.071	0.590	73	138
	2.817	2.812	1.823	10	179
	2.955	2.953	2.016	9	170
3	1.208	1.209	0.600	85	106
	1.567	1.567	0.606	104	105
	1.648	1.644	1.860	22	113
	1.762	1.761	2.057	19	106
4	0.985	0.985	0.624	123	56
	1.645	1.645	0.629	114	87
	1.267	1.264	1.904	36	52
	1.320	1.319	2.106	32	49
5	1.129	1.129	0.655	112	60
	1.935	1.935	0.655	95	116
	1.474	1.473	1.952	59	32
	1.462	1.461	2.163	55	33
6	1.408	1.407	0.690	80	100
	2.247	2.247	0.668	78	142
	1.782	1.782	1.997	74	38
	1.737	1.737	2.222	65	35
7	2.252	2.247	0.743	61	155
	2.643	2.635	0.718	60	166
	2.108	2.105	2.050	68	64
	2.267	2.348	2.328	65	61
8	0.987	0.987	0.532	42	135
	0.791	0.793	0.261	62	141
	1.280	1.277	1.538	2	146
	1.391	1.392	1.756	2	132
9	0.632	0.632	0.545	41	94
	0.505	0.506	0.261	75	109
	0.753	0.749	1.555	0	70
	0.897	0.898	1.779	2	63
10	0.333	0.333	0.561	35	42
	0.014	0.008	0.265	0	0
	0.326	0.315	1.578	2	3
	0.563	0.564	1.806	7	11

Table 5 (Cont'd.)
The relationship between the actual ISDs and the Predictive Densities

Option	Root-Mean Sq. Error - Mean	Root-Mean Sq. Error - Median	Average 90%ile band	# Sample Below band	# Sample Above band
11	0.259	0.258	0.581	44	16
	0.429	0.431	0.274	91	84
	0.599	0.595	1.605	34	0
	0.704	0.705	1.840	36	3
12	0.538	0.538	0.603	65	79
	0.707	0.709	0.290	87	117
	0.947	0.946	1.637	83	2
	0.985	0.986	1.878	70	5
13	0.991	0.991	0.479	98	92
	1.257	1.257	0.544	102	98
	1.007	1.008	1.369	54	46
	1.007	1.006	1.625	51	34
14	0.837	0.837	0.489	99	79
	1.119	1.120	0.534	104	93
	0.952	0.952	1.380	83	24
	0.982	0.981	1.642	79	16
15	0.714	0.714	0.501	89	73
	0.986	0.987	0.535	100	86
	1.014	1.011	1.463	119	9
	1.052	1.050	1.666	103	7

Notes: This table reports aspects of the posterior densities of the model parameters and functions of interest. The model is estimated using weekly data from 1989 through October 1993, on four different sets of options and the underlying asset (S&P 500 index). The fours sets are: 1) All 15 options; 2) Options 8, 10, and 12; 3) Option 10; 4) Option 10, with $\Sigma_{(1,1)} = 0.2$.
The option characteristics are presented in Table 2.
The model we consider for the evolution of stock returns is:

$$dS = \mu S dt + \sqrt{v} S dz$$

$$dv = \kappa(\theta - v)dt + \sigma\sqrt{v}d\omega$$

The model for option prices is: $P(t, X, T) = C(t, v_t, X, T, \cdot) + \epsilon$.
$\epsilon \sim N(0, \Sigma)$. C is Heston's(1993) model of option pricing with stochastic volatility

As explained in the text, identification is achieved by placing a diagonal prior on Σ.
The posterior density of the implied standard deviation is constructed using the 100,000 Gibbs draws for each of the 249 dates in the sample. The percentile bands are constructed using the actual option constructed using the 100,000 Gibbs draws for each of the 249 dates in the sample. The percentile bands are constructed using the actual option characteristics along with the Gibbs draws on the volatility and parameters for each date. The variance-covariance matrix (Σ) is not used in this context. The "# Above band" shows the number of the 249 days in the sample when the actual isd was above the 90%ile posterior band on that date.

Table 6
The Slope of the Smile and Term

Smile Measure:	Option 2 Option 6	Option 8 Option 12	Option 9 Option 11	Option 15 Option 13
Actual Data				
Group 1	0.022069792	0.023097934	0.012363836	0.010716495
Group 2	0.076942917	0.058587693	0.030519542	0.026160277
Group 3	0.085942970	0.059424316	0.030333107	0.023473363
Group 4	0.064307309	0.040146356	0.020320922	0.014571154
Posterior Function				
Group 1				
Mean	0.0282427	0.0252901	0.0125905	0.0114205
Std. Dev.	0.0004271	0.0003214	0.0001555	0.0001195
10%ile	0.0276885	0.0248738	0.0123890	0.0112647
Median	0.0282464	0.0252950	0.0125926	0.0114230
90%ile	0.0287883	0.0257015	0.0127895	0.0115730
Group 2				
Mean	0.0594924	0.0530041	0.0265577	0.0264453
Std. Dev.	0.0008726	0.0006527	0.0003172	0.0002752
10%ile	0.0583654	0.0521603	0.0261470	0.0260878
Median	0.0594991	0.0530125	0.0265619	0.0264502
90%ile	0.0606057	0.0538358	0.0269615	0.0267967
Group 3				
Mean	0.0474226	0.0431046	0.0215845	0.0215052
Std. Dev.	0.0006942	0.0005415	0.0002650	0.0002242
10%ile	0.0465256	0.0424041	0.0212416	0.0212141
Median	0.0474276	0.0431124	0.0215885	0.0215090
90%ile	0.0483077	0.0437959	0.0219220	0.0217911
Group 4				
Mean	0.0248525	0.0231032	0.0115523	0.0110299
Std. Dev.	0.0003609	0.0002993	0.0001471	0.0001171
10%ile	0.0243868	0.0227165	0.0113622	0.0108778
Median	0.0248557	0.0231070	0.0115541	0.0110318
90%ile	0.0253133	0.0234852	0.0117400	0.0111792

Notes: This table reports the average slope of the implied volatility smile for different option groups. The 249 observations are sorted into four groups based on the short-term isd slope (i.e., the difference between the actual isd on Option 1–the shortest term, deepest in-the-money option, and Option 7–the shortest term, deepest out-of-the-money option). The model is estimated using the panel of 15 options.

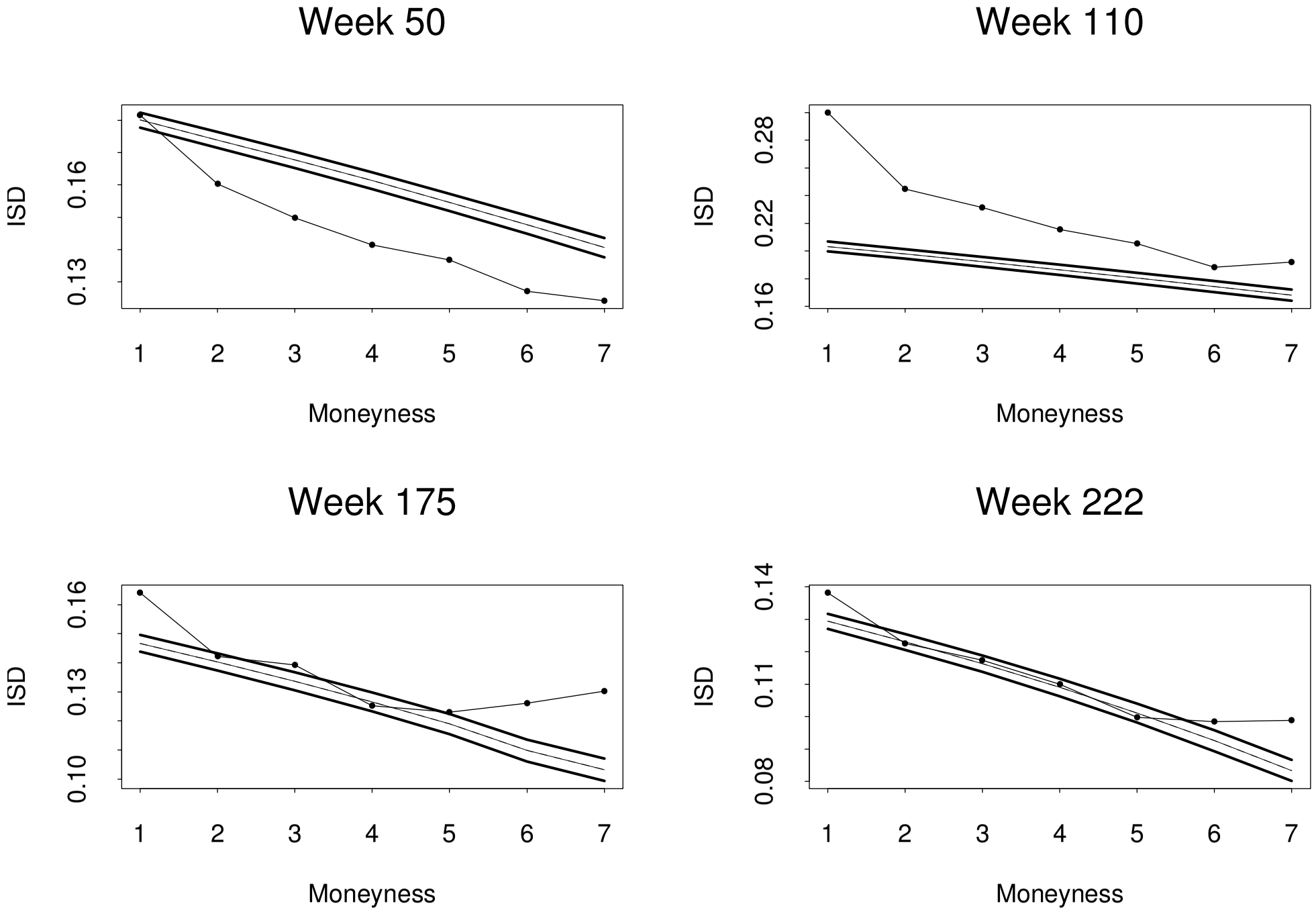

Figure 1 Predictive Densities and Actual Smiles –– Nearest–Term Options.

Actual implied volatilities are represented as dots. The 90th and 10th percentiles of the posterior predictive density are plotted around the median of the predictive density for each of the seven near–term options on each of the four dates.

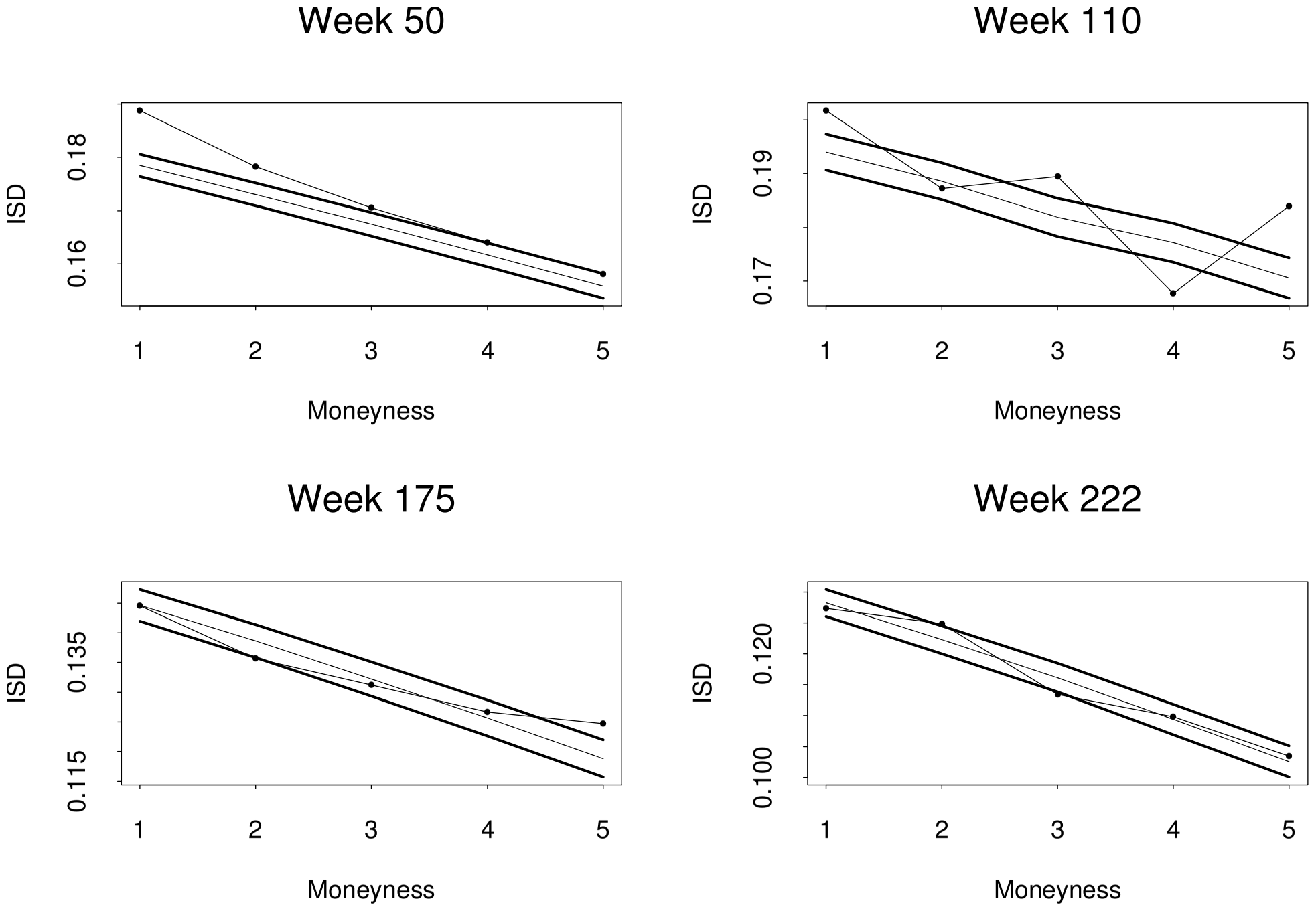

Figure 2 Predictive Densities and Actual Smiles –– Intermediate–Term Options.

Plotted: 95%ile, Median and 25%ile from Predictive Density, and Actual Smile (dotted).

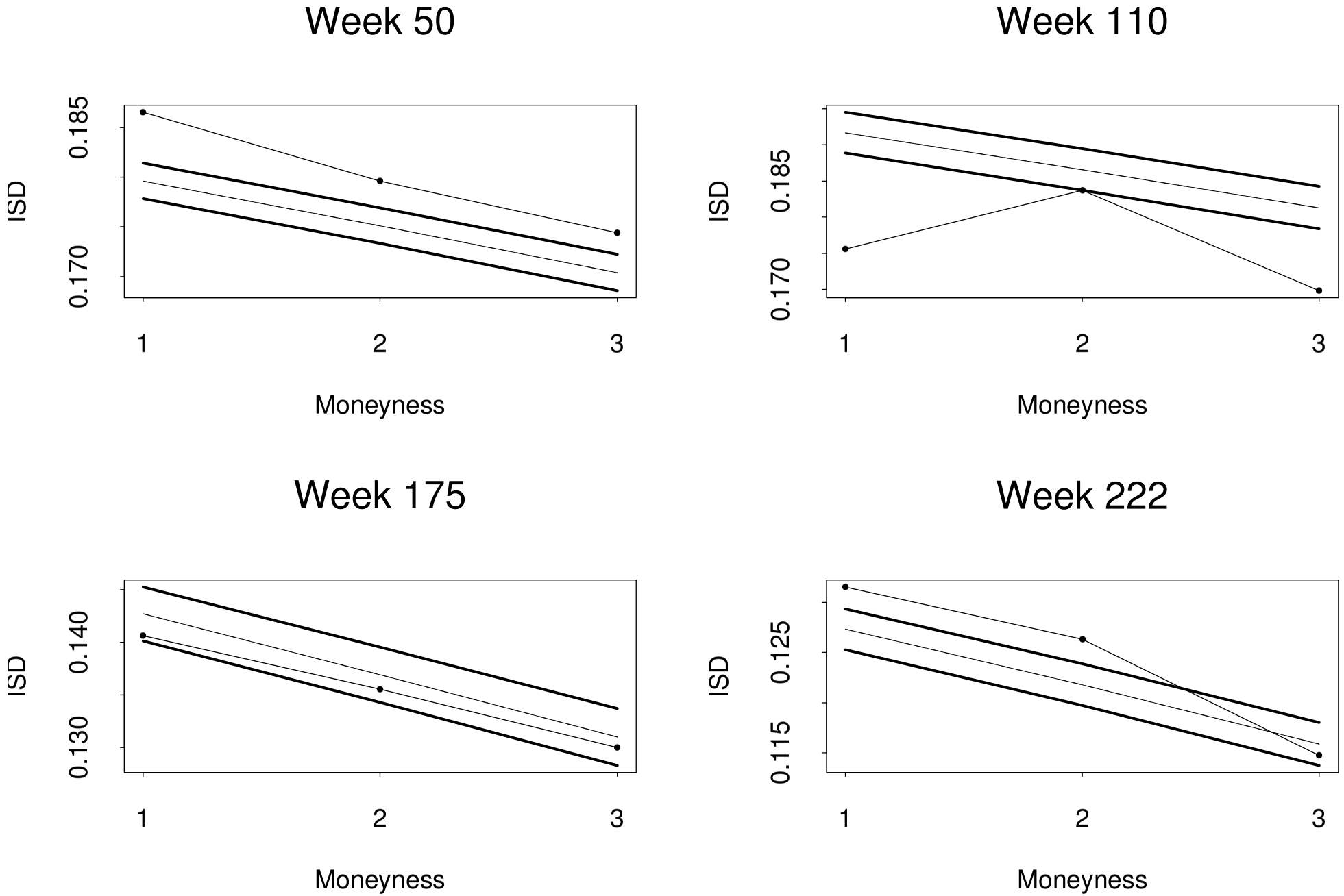

Figure 3 Predictive Densities and Actual Smiles –– Longest–Term Options.

Plotted: 95%ile, Median and 25%ile from Predictive Density, and Actual Smile (dotted).

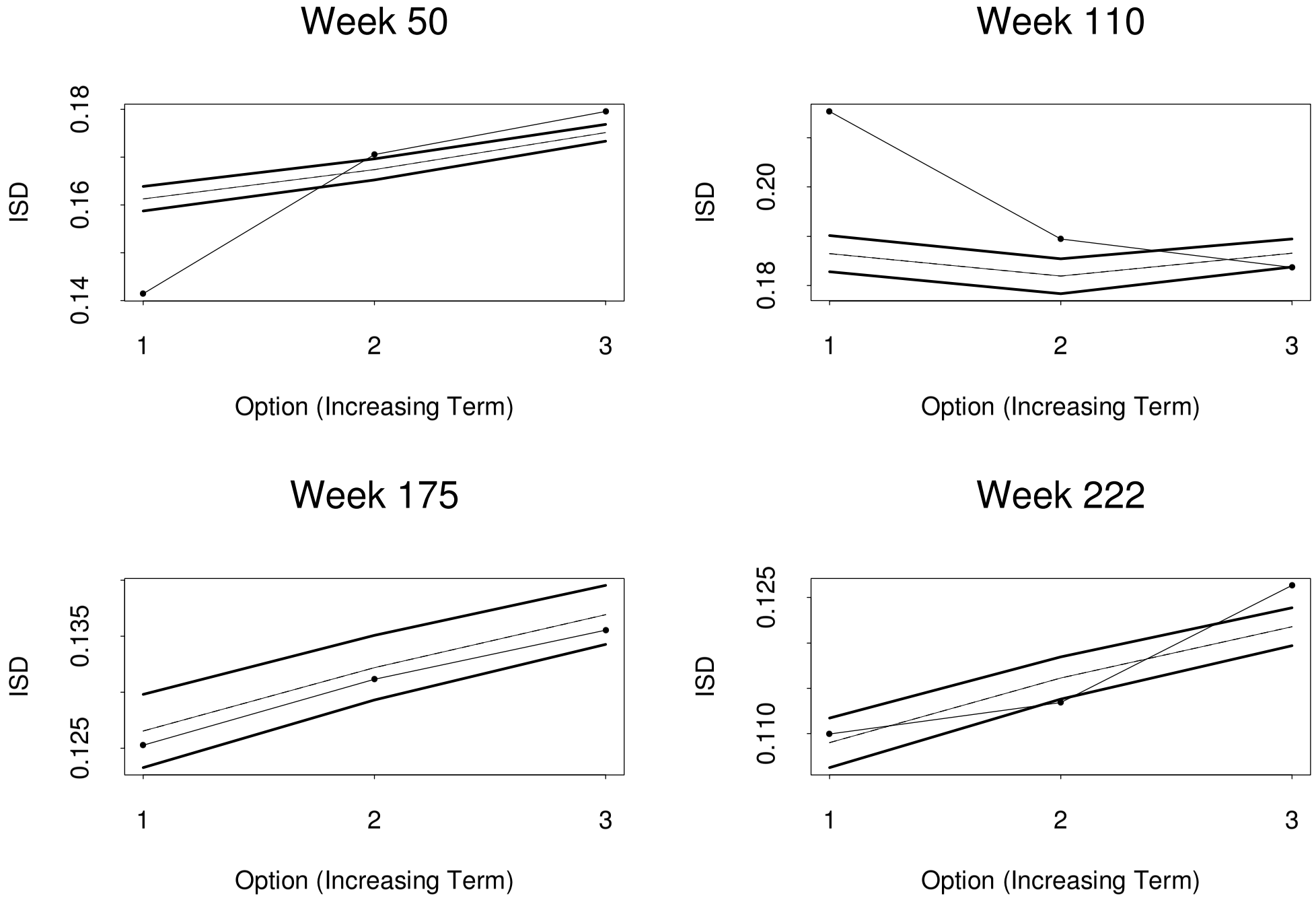

Figure 4 Predictive Densities and Actual Implied Volatility Term Structures.

Actual implied volatilities are represented as dots. The 95th and 5th percentiles of the posterior predictive density from the Heston model are plotted around the median of this predictive density for each of the three at–the–money options on each of the four dates.